Natural Resources and Spatial Spillovers

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Dulguun Batbold

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

TERRY ROE

June, 2015

Acknowledgements

This dissertation would not have been possible without the generous guidance and advice of my committee members as well as love and support from my family.

Chief among the people I owe my gratitude is my advisor, friend, and mentor, Professor Terry Roe. I cannot thank him more for his unwavering support, insightful advice, and kind mentorship. My thanks also go to my committee members Timothy Kehoe, Rodney Smith, and Gerard McCullough, for their deeply probing and helpful comments and suggestions as well as their encouragement.

I would also like to extend my thanks to my colleague and friend at the Federal Reserve Bank of Minneapolis Rob Grunewald. This thesis is a natural outgrowth of the many conversations we've shared and *fedgazette* articles we've coauthored on the shale energy boom in the Bakken area.

Last, I would like to thank my family for their love, encouragement, and support.

Dedication

To my parents who raised me to be the man I am, to my wife who adds meaning and beauty to my life, and to my son Erdem who is my pride and hope

Abstract

Regions going through a natural resource boom tend to have higher average incomes and employment relative to the rest of the country. For policy analysis, a question that often needs to be answered is to what extent the economic growth in the extraction region spills over to neighboring areas. This thesis develops a detailed methodology for analyzing the economic effects of geographically localized shocks within the framework of a parsimonious spatial general equilibrium model, including various methods for estimating key parameters.

This model-based approach is being offered as a complementary tool for applied researchers conducting economic impact analysis. Existing empirical methods such as input-output analysis or difference-in-difference estimation techniques are often not optimal for analyzing spatially correlated data, and this model-based methodology can be used to overcome their limitations. Another important advantage of this methodology is that it is computationally tractable and has a relatively low data requirement, which can make a particularly big difference in studying developing countries where data quality and availability can often be an insurmountable challenge.

Following the exposition of the methodology, this thesis presents two separate applications, one involving a developed nation and the other a developing one. In the first case, the methodology is applied to analyze the economic impact of the shale energy boom that's been occurring in and around Bakken counties in western North Dakota and eastern Montana over the past decade. In the second case, the methodology is used to analyze the economic impact of the Oyu Tolgoi copper-gold mining project in the Southern Gobi region of Mongolia.

A common conclusion that is drawn from the two applications mentioned above is that economic booms fueled by natural resource extracting industries are largely local and have limited spillover effects on neighboring regions.

Contents

Acknowledgements **i**

Dedication **ii**

Abstract **iii**

List of Tables **vi**

List of Figures **vii**

1 Introduction **1**

2 Spatial Equilibrium Model **7**

2.1 Base Model 7

2.2 Extended Model 16

3 Bakken Shale and Spatial Spillovers **24**

3.1 Shale Energy Revolution 24

3.2 Bakken Impact 40

4 Economic Impact of the Oyu Tolgoi Project **53**

4.1 Oyu Tolgoi Project 55

4.2 Oyu Tolgoi Economic Impact 61

5 Conclusion and Discussion **66**

References **68**

Appendix A. Two-Region Example **72**

List of Tables

3.1 Employment impact of the Bakken oil boom 49
3.2 Wage impact of the Bakken oil boom 50
4.1 Employment impact of the Oyu Tolgoi project 63
4.2 Wage impact of the Oyu Tolgoi project 64
4.3 Sensitivity analysis of the employment effect 65
4.4 Sensitivity analysis of the wage effect 65

List of Figures

1.1 Labor migration to Bakken counties by origin 4
1.2 Interstate shipment shares by distance to source and destination 5
2.1 Basic model: spatial economy on a line 14
2.2 Basic model: spatial impulse responses 15
2.3 Real wages estimates based on BEA relative price parity data 16
2.4 Extendend model: spatial impulse responses 22
2.5 Spatial impulse responses of wages with and without congestion effects . 23
3.1 Bakken and Three Forks Formations, U.S. Geological Survey 25
3.2 Dry hole percentage of crude oil and natural gas wells drilled 26
3.3 North America rig count by drilling type, 1991-2013 27
3.4 Monthly U.S. crude oil production, 2000-2014 29
3.5 Decline curves for the Bakken, 2004-2014 vintage wells 30
3.6 U.S. county average wage map pre- and post-Bakken boom 32
3.7 U.S. county wage growth map, 2003Q1 to 2014Q2 33
3.8 U.S. county employment map pre- and post-Bakken boom 34
3.9 U.S. county employment growth map, 2003Q1 to 2014Q2 35
3.10 Moran's I plots of employment and wage growth 36
3.11 100-mile band rings around the Bakken 37
3.12 Employment and wage indices by distance to the Bakken 38
3.13 Commodity shipments to and from North Dakota by source and destination 41
3.14 Estimating Armington elasticity 42
3.15 Calibrating the congestion parameter 43
3.16 Calibrating share of household expenditures on non-traded goods 45
3.17 Utility weight estimates 47

3.18 Estimated impact of the Bakken oil boom relative to counterfactual paths 49
3.19 Estimated productivities by county 51
3.20 U.S. county productivity growth map 52
4.1 Employment map of Mongolia 56
4.2 Aimags by distance to the Oyu Tolgoi mine 59
4.3 Employment and wage indices by distance to Oyu Tolgoi mine 60
4.4 Wage growth map by aimags, 2007Q1 to 2013Q1 61

Chapter 1

Introduction

Regions going through a natural resource boom are typically characterized by a fast pace of growth in both employment and wages relative to their peers in other parts of the country. For policy analysis, there is often a strong interest in obtaining estimates of the extent to which the economic growth in the extraction region can be expected to spill over to neighboring areas. For one thing, uneven growth can have important distributional implications for policymakers in regional governments, who may be interested in knowing these estimates when allocating infrastructure investment budgets or planning urban development. For another thing, the extent and strength of such spillovers can be important factors to consider for governments doing cost-benefit analysis of large investments into a mining project, especially for developing nations like Mongolia for whom the mining sector constitutes an important source of government revenues as well as a strategic means of promoting regional development.

This thesis develops a methodology for analyzing the geographic pattern of economic impact from localized shocks within the framework of a spatial general equilibrium model. The choice of this model-centric approach is motivated by two considerations. The first is that existing empirical methods have strong limitations when applied to spatially correlated data. A common method for estimating economic impact of commodity booms, for instance, is to employ multipliers from regional industry input-output tables. A report by Tunstall et al (2014) exemplifies this type of analysis, in which the authors use input-output table coefficients to derive economic impact estimates of the oil production from the Eagle Ford Shale in Texas.

Estimates based on this type of analysis, however, generally assume a stable input-output relationship between industries, which may be problematic for analyzing booming regions, which typically undergo a drastic structural change of their economy. The Bakken area in North Dakota, for instance, is one example, which has gone through a very rapid transition from a largely rural economy dominated by agriculture to an energy-based one dominated by the oil industry as a result of the shale energy revolution. Grunewald and Batbold (2013) documented how the mining and oil extraction industry accounted for most of the phenomenal growth in North Dakota's taxable sales since the start of the shale oil boom, much of it accounted for by counties in the immediate neighborhood of the oil producing area. More problematically, input-output tables assume stable relative prices between industries, which can be hard to maintain in cases like the Bakken, given the large and growing disparities in average wages of oil and non-oil industries in the state.

Another commonly utilized method of analyzing economic impact is the difference-in-differences (DID) estimation. Alcott and Keniston (2013), for example, employ the DID method to estimate economic impact from booming oil and gas activity in the U.S. and conclude that it has increased growth rates in producer counties by 60 to 80 percent relative to non-producer counties. A key limitation of such difference-in-difference types of estimations is the possibility of spatial spillovers from producing regions to neighboring non-producing regions, which, if exist, can bias differential impact estimates.

The second consideration that motivated the choice of the model-based methodology outlined in this thesis is that it is both computationally tractable and also has a relatively low data requirement. The latter especially is a big advantage for researchers of developing countries like Mongolia, where data quality and availability can often be an insurmountable challenge. Input-output tables in particular are very data-intensive, which is why regional input-output tables are hard to find for developing countries and, even if found, are often outdated.

Following the exposition of the model, two applications of the methodology are discussed in this thesis. First, the model is applied to analyze spatial pattern of the economic impact from the shale oil boom that has recently swept through counties in western North Dakota and eastern Montana. The other case covers Mongolia, a small developing nation that has been undergoing a major mining boom, following the discovery of a large copper-gold deposit named Oyu Tolgoi in the Southern Gobi region of the country. As will be documented later in the thesis, in both cases there can be observed a spatial pattern to the impact with the wage and employment growth appearing strongest in the core mining areas and dissipating with the distance away from the epicenter of mining activity.

There are two main channels explored in this thesis through which a productivity shock in a given location may theoretically result in such a pattern of spatial spillovers. One possibility is the labor migration channel, which can serve as a means through which a localized impulse can be spatially transmitted. To the extent that inter-regional labor mobility may be a function of distances between them, this labor reallocation channel can potentially account for the observed dissipating spatial pattern.

Goods trade between regions is the other explored channel which can potentially explain the observed pattern. The methodology developed in this thesis primarily focuses on this second channel. For one thing, existing theoretical models of spatial distribution of economic activity and employment for the most part have little to say about the geographical distribution of labor flows. If labor allocation across regions is assumed to be endogenous, spatial equilibrium typically requires workers to be indifferent between locations, which helps pin down the equilibrium distribution of workers among different regions but at the same time serves as a source of indeterminacy in terms of labor flows needed for the reallocation of labor to occur from one observation to the next.

For another thing, in the case of the Bakken shale oil boom, the available data show very little labor migration. Looking at the county-to-county worker flow data from the Internal Revenue Service (IRS), for instance, one can make two observations. First, despite the phenomenal wage differential between the Bakken region and some of the

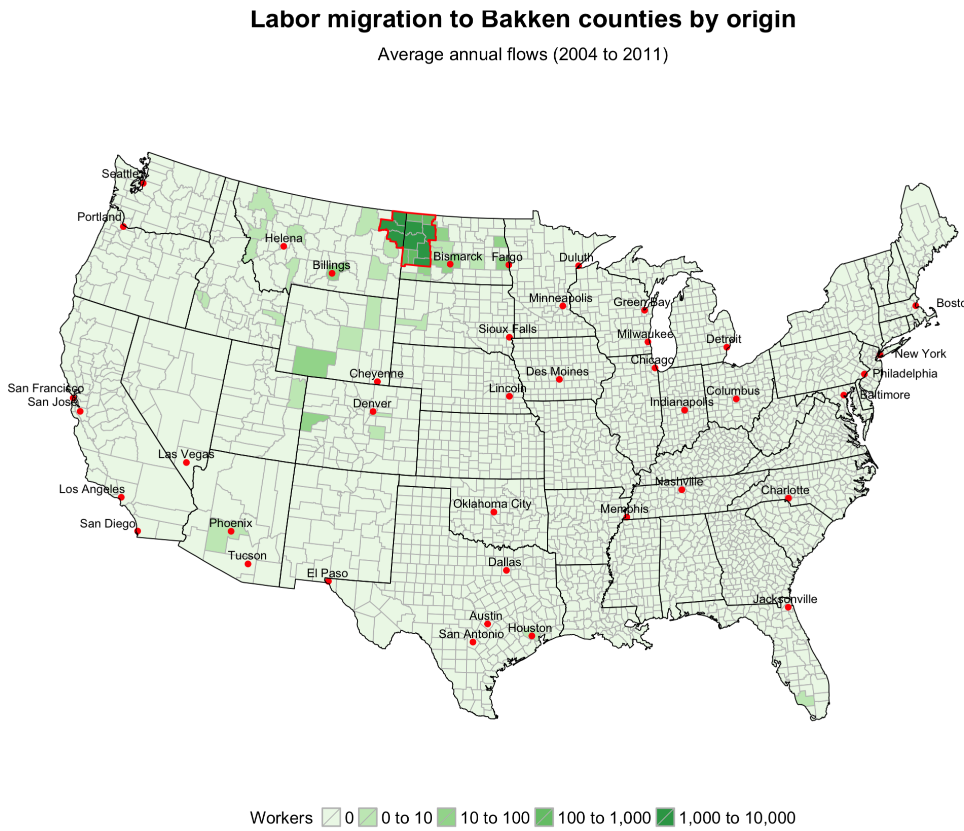


Figure 1.1: Based on the Internal Revenue Service's County-to-County Migration Data

neighboring counties, there has been a relatively modest outflow of workers to the region between 2004 and 2011. The overwhelming majority of counties even within North Dakota and Montana had no labor movement into the oil-producing counties at all during this period, according to the IRS data. Second, even for counties that did lose workers to the Bakken region, the number of workers moved constituted a tiny fraction of total workforce in originating counties, with virtually all of them contributing less than 0.3 percent of workers to the oil boom region. This suggests, that the observed spatial pattern of economic impact is unlikely to be accounted for solely by labor movements into these booming counties.

On the other hand, if we look at the Commodity Flow Survey data by the U.S. Department of Transportation, interstate shipment volumes likewise appear spatially clustered, with shipments originating from any given state inversely related to the geographical distance to destination states and positively related to the size of those states, very much consistent with the gravity models of trade.

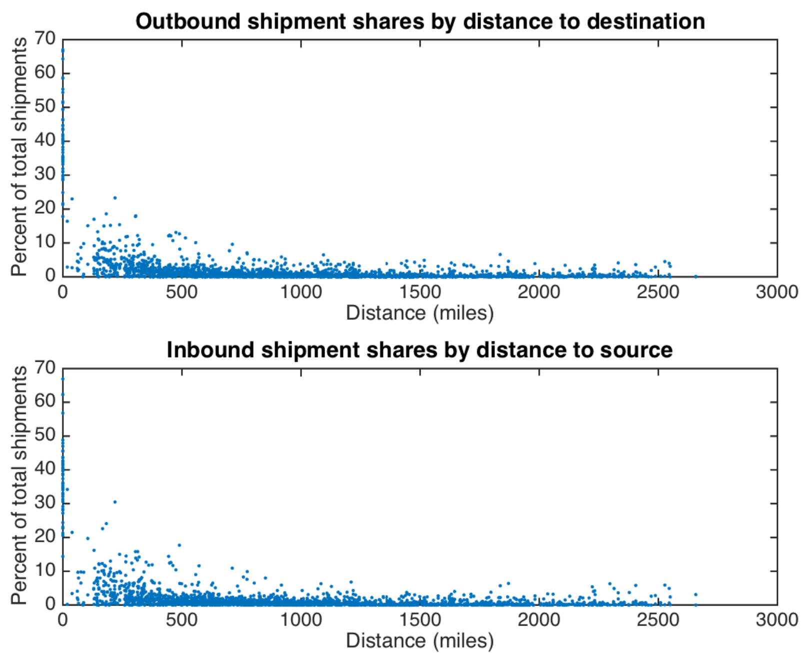


Figure 1.2: Based on the Department of Transportation's 2007 Commodity Flow Survey Data

While there are other possible channels such as the commodity boom's impact on government revenues or direct productivity spillovers, there aren't as compelling reasons to believe that these channels will generate the observed spatial pattern of impact. Though they are important, a more in-depth exploration of these channels is left for future research.

The rest of the thesis is organized as follows:

· Chaper 2 introduces the spatial general equilibrium model underpinning the analysis and outlines a procedure through which the model can be used to impute regional productivity levels from employment and wage data and construct counterfactual paths for economic impact analysis.

· Chapter 3 of the thesis examines the impact pattern of the shale energy revolution on Bakken counties and applies the methodology to derive estimates of economic spillovers on the surrounding regions.

· Chapter 4 presents another application of the methodology, used to analyze economic impact of the Oyu Tolgoi project in Mongolia.

· Chapter 5 presents a final discussion of the analyses presented in the thesis.

Chapter 2

Spatial Equilibrium Model

In this chapter, I will outline a simple version of a spatial general equilibrium model, which will form the framework for the empirical analysis of the economic impact of the Bakken shale energy boom in Chapter 3 as well as the Oyu Tolgoi mining project in Chapter 4.

The advantages of having a fully-specified model of a spatial economy is that it enables modeling of counterfactual paths for variables of interest, which would allow structural estimation of the effects of localized productivity shocks as well as analysis of their spatial propagation patterns.

2.1 Base Model

Model environment

Consider an Armington (1969) style setup where a set S of geographical regions each produce a differentiated good using labor as the only input. In the base case, assume regions differ only in terms of their exogenous productivities A_i and their geographical locations, as expressed by the iceberg transportation costs of moving goods from location s to location i, denoted by $\tau_{si} > 1$ and $\tau_{ii} = 1$ for all $s, i \in S$.

Households inelastically supply their labor in each location i and optimally allocate their labor income on goods from different regions to maximize their utility:

$$W_i = \left[\sum_{s \in S} a_s y_{si}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{2.1}$$

where a_s are utility weights on goods from region s and y_{si} are goods from region s consumed by households in region i. As standard, optimality conditions imply region i's household expenditure on goods from location s takes the following form:

$$p_{si} y_{si} = Y_i \cdot P_i^{\sigma} \cdot \left[a_s^{\sigma} p_{si}^{1-\sigma} \right] \tag{2.2}$$

Here Y_i and P_i are region i's composite consumption and price index, respectively. Household's budget constraint then implies that its aggregate consumption (and utility) is equal to the real wage:

$$Y_i = \frac{w_i}{P_i} = W_i \tag{2.3}$$

where the region's price index is given by:

$$P_i = \left[\sum_{s \in S} a_s^{\sigma} p_{si}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \tag{2.4}$$

The household's budget constraint would then imply that the share of income spent by consumers in region i on goods from region s will be proportional to the weight of the relevant price in the regional price index:

$$\pi_{si} = \frac{p_{si} y_{si}}{w_i} = \frac{a_s^{\sigma} p_{si}^{1-\sigma}}{P_i^{1-\sigma}} \tag{2.5}$$

Firms in any given region s are assumed to be competitive and optimize supply of that region's good to maximize profits:

$$\begin{aligned} \max_{y_{si}^f,\ \ell_s^f} \quad & \sum_i p_{si} y_{si}^f - w_s \ell_s^f \\ \text{s.t.:} \quad & \sum_i \tau_{si} y_{si}^f \leq A_s \ell_s^f \end{aligned}$$

Competitive markets imply firms set prices at marginal cost and earn zero profits:

$$p_{si} = \tau_{si} \cdot \frac{w_s}{A_s} \tag{2.6}$$

Goods markets and local labor markets must clear, which means that for each region s total shipments must equal total production:

$$\sum_i \tau_{si} y_{si} L_i = A_s L_s \tag{2.7}$$

Equivalently, market clearing condition (2.7) can be rewritten in terms of expenditures by multiplying both sides by w_s/A_s and using equations (2.5) and (2.6) as follows:

$$\begin{aligned} \frac{w_s}{A_s} \cdot A_s \cdot L_s = \sum_i \tau_{si} \cdot \frac{w_s}{A_s} \cdot y_{si} \cdot L_i \; &= \; \sum_i p_{si} \cdot y_{si} \cdot L_i \\ w_s L_s = \sum_i \left[\frac{a_s^\sigma p_{si}^{1-\sigma}}{P_i^{1-\sigma}} \right] \cdot w_i L_i \; &= \; \sum_i \pi_{si} \cdot w_i L_i \end{aligned} \tag{2.8}$$

so that total income of region s equals to the sum of expenditures by other regions on goods from region s. Labor markets at the economy-level must also clear so that aggregate labor demand equals aggregate labor supply:

$$\sum_{s \in S} L_s = \bar{L} \tag{2.9}$$

In a spatial equilibrium with mobile labor, workers must have no incentive to relocate and thus be indifferent between locations, i.e. $W_i = W_j$ for all $i, j \in S$, which in turn implies real wages are equalized across regions:

$$\frac{w_i}{P_i} = \frac{w_j}{P_j} \qquad \text{for all} \quad i, j \in S \tag{2.10}$$

Equilibrium

Equilibrium of this model is defined as prices $\{w_s,\ p_{si}\}_{s,i \in S}$ and allocations $\{L_s,\ y_{si}\}_{s,i \in S}$ that are consistent with the optimizing behavior of households and firms and satisfy market clearing and welfare equalization conditions, given an exogenous productivity vector A and the distance matrix τ, i.e satisfy equations (2.1) through (2.10).

Welfare equalization condition (2.10) implies that relative wages must equal relative prices:

$$W = \frac{w_i}{P_i} = \frac{w_j}{P_j} \qquad \Rightarrow \qquad w_i \;=\; W \cdot P_i \;=\; W \cdot \left[\sum_s a_s^{\sigma} p_{si}^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \tag{2.11}$$

Although the price index P_i on the righthand side of (2.11) is a nonlinear function of wages, we can substitute out prices using (2.6) to obtain the following system:

$$w_i^{1-\sigma} = W^{1-\sigma} \cdot \sum_s a_s^{\sigma} \cdot \tau_{si}^{1-\sigma} \cdot A_s^{\sigma-1} \cdot w_s^{1-\sigma} \tag{2.12}$$

which is now linear in $w^{1-\sigma}$. Here W is a scalar equal to the equilibrated welfare (real wage) level attained across regions. We can rewrite system (2.12) in matrix form:

$$\mathbf{M}\mathbf{w}^{1-\sigma} = W^{\sigma-1} \cdot \mathbf{w}^{1-\sigma} \tag{2.13}$$

where each element of the square matrix $\mathbf{M}$ is defined by $m_{ij} = a_j^{\sigma} \cdot \tau_{ji}^{1-\sigma} \cdot A_j^{\sigma-1}$. Then, given an exogenous matrix $\mathbf{M}$, transformation of equilibrium wages $\mathbf{w}^{1-\sigma}$ can be solved as the eigenvector of $\mathbf{M}$, with the associated largest eigenvalue $W^{\sigma-1}$.

Note that $\mathbf{M}$ is a positive and square matrix. Correspondingly, by the Perron-Frobenius Theorem, there exists a unique (to a scale) strictly positive eigenvector $\mathbf{v} = \mathbf{w}^{1-\sigma}$ of $\mathbf{M}$ such that all elements of $\mathbf{w}^{1-\sigma}$ are real and strictly positive, so the existence and uniqueness of equilibrium relative wages are assured.

Note also that uniform scaling of nominal wages doesn't change the average welfare, which can only grow with increases in productivity. By property of eigenvalues, any scaling $\gamma\mathbf{M}$ will result in a proportionately scaled eigenvalue $\gamma W^{\sigma-1}$. Therefore, we can conclude that wages *everywhere* respond to aggregate productivity shocks much more than prices, consistent with (2.6) where we know:

$$\frac{\partial(w_s/p_{si})}{\partial A_s} = \frac{1}{\tau_{si}} > 0$$

With wages known, prices can be readily identified (to a scale) using (2.6). Expenditure shares can then be computed using equation (2.5):

$$\pi_{si} = \frac{a_s^\sigma \tau_{si}^{1-\sigma} w_s^{1-\sigma} A_s^{\sigma-1}}{P_i^{1-\sigma}} \;=\; a_s^\sigma \tau_{si}^{1-\sigma} A_s^{\sigma-1} \cdot \frac{w_s^{1-\sigma}}{P_s^{1-\sigma}} \cdot \frac{P_s^{1-\sigma}}{P_i^{1-\sigma}}$$

$$\pi_{si} = W^{1-\sigma} \cdot a_s^\sigma \tau_{si}^{1-\sigma} A_s^{\sigma-1} \cdot \left[\frac{w_s}{w_i}\right]^{1-\sigma} \tag{2.14}$$

While the equilibrium allocations are homogeneous of degree zero in wages and prices, the ratio of wages to price indices is strictly governed by the scalar W contained in the eigenvalue. Using equation (2.14) then, we can rearrange (2.5) to solve for gross shipments (inclusive of the transportation cost) as follows:

$$y_{si} = \pi_{si} \cdot \frac{w_i}{p_{si}} \;=\; \pi_{si} \cdot \frac{w_i}{\tau_{si} w_s} \cdot A_s \;=\; \frac{1}{\tau_{si}} \cdot W^{1-\sigma} \cdot a_s^\sigma \tau_{si}^{1-\sigma} A_s^{\sigma-1} \cdot \left[\frac{w_s}{w_i}\right]^{1-\sigma} \cdot \frac{w_i}{w_s} \cdot A_s$$

$$\tau_{si} y_{si} = W^{1-\sigma} \cdot a_s^\sigma \tau_{si}^{1-\sigma} \cdot \left[\frac{w_i}{w_s}\right]^{\sigma} \cdot A_s^\sigma \tag{2.15}$$

Then, the market clearing condition (2.7) will be a linear system in L:

$$A_s L_s = \sum_i \tau_{si} y_{si} \cdot L_i = W^{1-\sigma} \cdot \sum_i a_s^\sigma \tau_{si}^{1-\sigma} \cdot \left[\frac{w_i}{w_s}\right]^{\sigma} \cdot A_s^\sigma \cdot L_i$$

$$W^{\sigma-1} L_s = \sum_i a_s^\sigma \tau_{si}^{1-\sigma} \cdot \left[\frac{w_i}{w_s}\right]^{\sigma} \cdot A_s^{\sigma-1} \cdot L_i \tag{2.16}$$

which in matrix form can be written as:

$$\mathbf{HL} = W^{\sigma-1}\mathbf{L} \tag{2.17}$$

where each element of matrix $\mathbf{H}$ is given by $h_{si} = a_s^\sigma \tau_{si}^{1-\sigma} \cdot w_i^\sigma / w_s^\sigma \cdot A_s^{\sigma-1} > 0$. By the Perron-Frobenius Theorem again, $\mathbf{L}$ will be the unique (to a scale) positive and real-valued eigenvector of $\mathbf{H}$ and $W^{\sigma-1}$ will be the associated (largest) eigenvalue. The aggregate labor market clearing condition (2.9) can then be used to scale the equilibrium labor allocations.

Imputation of local productivities from wage and employment data

In applied uses of the model, one might wish to invert the process and impute the model's implied exogenous productivities from the observed labor market data. If the model's parameters (a_i, σ, τ_{ij}) are known, observed wages and employment data by region are sufficient to perform this imputation, up to a constant.

If the observed data is assumed to represent an equilibrium outcome of the model, productivities must satisfy (2.8), which can be rewritten using (2.6):

$$\begin{aligned} w_s L_s &= \sum_i \left[\frac{a_s^\sigma \cdot \tau_{si}^{1-\sigma} \cdot w_s^{1-\sigma} \cdot A_s^{\sigma-1}}{P_i^{1-\sigma}} \right] w_i L_i \\ A_s^{1-\sigma} &= \sum_i a_s^\sigma \tau_{si}^{1-\sigma} \left[\frac{w_s}{P_s} \cdot \frac{P_s}{P_i} \right]^{1-\sigma} \frac{w_i L_i}{w_s L_s} \\ A_s^{1-\sigma} &= W^{1-\sigma} \cdot \sum_i a_s^\sigma \tau_{si}^{1-\sigma} \left[\frac{P_s}{P_i} \right]^{1-\sigma} \frac{w_i L_i}{w_s L_s} \quad = \quad \frac{W^{1-\sigma}}{w_s^\sigma L_s} \cdot \sum_i a_s^\sigma \tau_{si}^{1-\sigma} w_i^\sigma L_i \end{aligned}$$

The last expression uses the welfare equalization condition (2.10) which implies that $w_s/P_s = W$ for all regions $s \in S$. Equal real wages also imply that relative prices are equal to relative wages, i.e. $P_s/P_i = w_s/w_i$ for all pairs $s, i \in S$. Thus, relative productivities can be uniquely identified by:

$$\left[\frac{A_s}{A_t} \right]^{\sigma-1} = \frac{w_s^\sigma L_s}{w_t^\sigma L_t} \cdot \frac{\sum_i a_t^\sigma \tau_{ti}^{1-\sigma} w_i^\sigma L_i}{\sum_i a_s^\sigma \tau_{si}^{1-\sigma} w_i^\sigma L_i} \tag{2.18}$$

Given time series on county-level wages and employment, equation (2.18) then allows imputation of the relative productivity series. Note that larger regions (those with higher wages and larger populations) will generally be inferred to have higher relative productivities, while more geographically remote regions (those with larger τ) will, all else equal, be inferred to have lower relative productivities.

If one can find a reasonable way of disciplining the scale parameter for the productivity series, then counterfactual paths can be constructed to assess economic impact over time. One possible approach to scale this would be to match model's aggregate output $Y = \sum_s A_s L_s$ to its empirical counterparts such as the real GDP data.

Spatial impulse responses

To understand spatial patterns generated by the base model, consider a simple example of an economy on the line, which starts out with identical regions that have identical productivities and utility weights (i.e. $A_i = a_i = 1$ for all $i \in S$), as illustrated in Figure 2.1. Note that even in the symmetric case where all regions are identical, distribution of economic activity favors locations closer to the center of the line, where price indices are lower due to lower average distances to other locations on the line. Correspondingly, the model predicts central locations to have both higher employment and income relative to locations on the periphery.

If we let the relative productivity of the central region increase by increments of 1 percent, we can see the changes predicted by the model which shows a strong positive response in the economy of that region with employment and income rising and average prices falling.

In Figure 2.2 below are plotted changes to these variables, which notably feature a clear spatial pattern. The impulse response of price indices to the productivity shock is strongest at the epicenter and dissipates monotonically with distance from the source of the productivity shock. The region at the epicenter experiences strong growth in employment and income. Notably, regions farther away from the epicenter contribute more to the total flow of labor to the booming center. As a result, regions close to the epicenter experience an overall gain in the share of total employment and income while regions farther away from the source experience declines in their relative shares. It should be noted, however, that *all* regions gain from the productivity improvements as labor mobility equalizes real wages and welfare across locations at proportionately

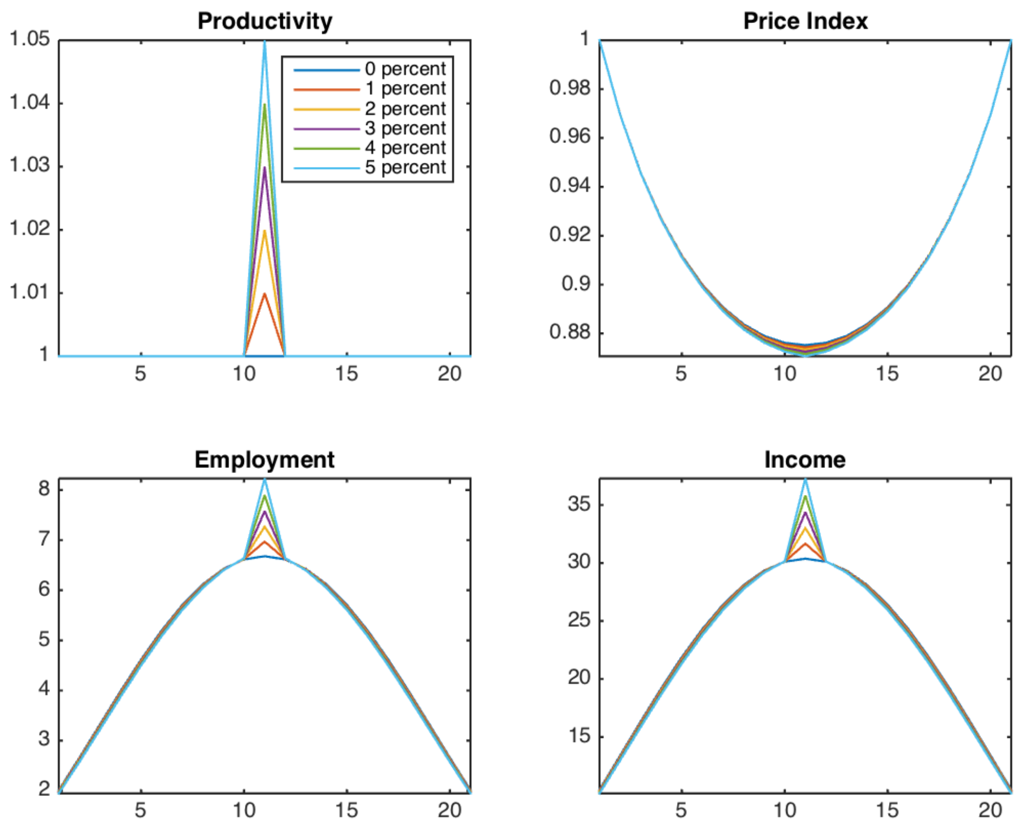


Figure 2.1: Symmetric economy on the line ($a_i = 1$ for all i, $\sigma = 5$)

higher levels.

While the base model is qualitatively able to generate the spatial patterns observed in connection with a mining boom, there are two important deficiencies inherent in the model that are strongly at odds with the data. First, by assumption the model maintains parity in real wages across locations, which is contradicted by the observed price parities. Moreover, this forced parity also counterfactually implies that nominal wages fall in the booming region even while they are rising in real terms. Second, the elasticity of employment response to productivity shocks is too large to be compatible with the relatively low interregional labor migration observed in the U.S. In the illustrative case of an economy on the line, for instance, a 1 percent increase in the relative productivity of the central region results in more than 4 percent increase in that region's employment.

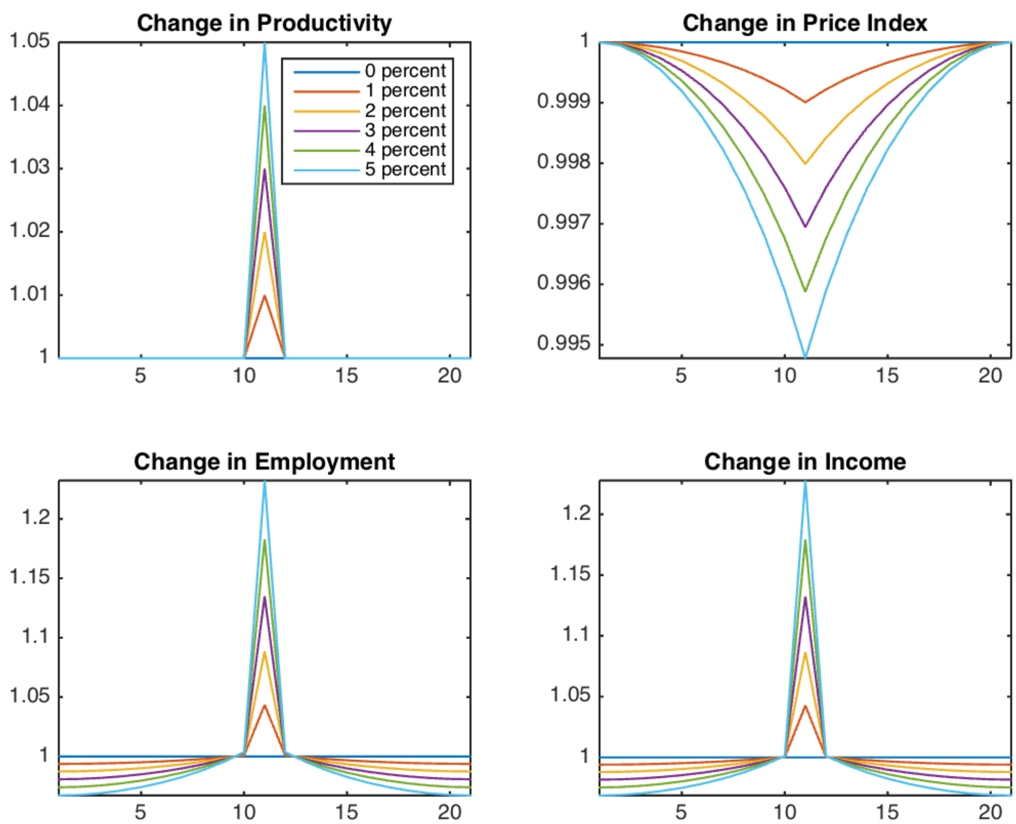


Figure 2.2: Spatial impulse responses of a symmetric economy on the line ($a_i = 1$ for all i, $\sigma = 5$)

Scaling the annual average weekly wage data from the Bureau of Labor Statistics (BLS) by the relative price parities reported by the Bureau of Economic Analysis (BEA), one can observe significant variation in thus estimated relative real wages. In 2008, for instance, real wages relative to that in North Dakota range from the low of 88 percent in Montana to the high of 170 percent in the District of Columbia. Moreover, estimated relative wages are not stable over time. While in 2008 39 states had estimated real wages higher than that in North Dakota, only 7 states still had higher estimated real wages than North Dakota in 2012. Montana's ranking in real wages has also risen from the last place in 2008 up two spots by 2012.

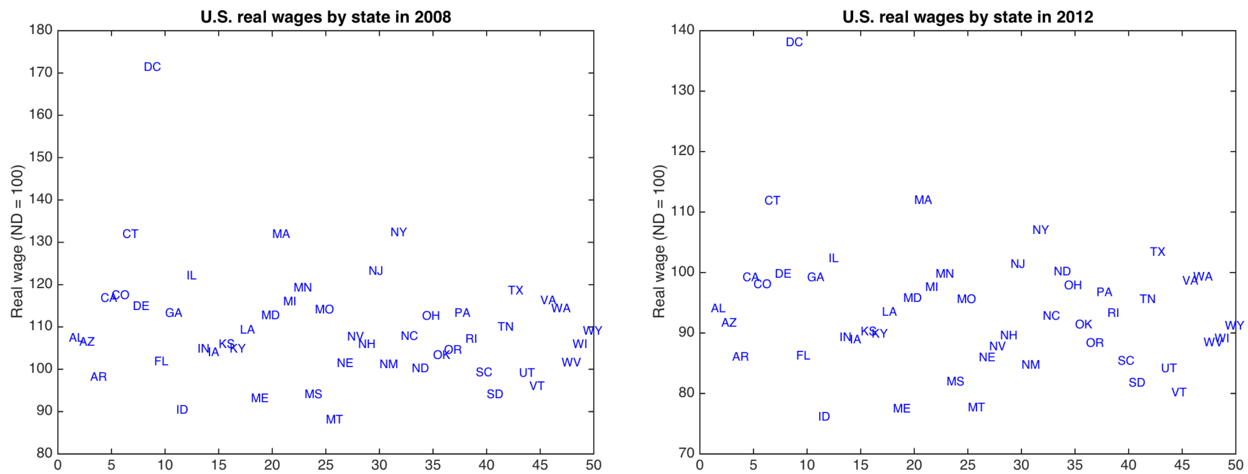


Figure 2.3: Real wage estimates are based on the Quarterly Census of Employment and Wages (QCEW) wage data and the BEA relative price parity data

Therefore, the base model needs modifications that would allow for variations in real wages across regions and reduce the elasticity of the employment response to productivity shocks to have a reasonable expectation of matching the data both qualitatively and quantitatively.

2.2 Extended Model

Extended model equilibrium

Consider now an extension of the base model as in Allen and Arkolakis (2014), featuring an additional source of heterogeneity across regions in terms of the amenities they offer to residents so that households' preferences are described by:

$$W_i = \left[\sum_{s \in S} a_s y_{si}^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} U_i \tag{2.19}$$

where U_i are composite local amenities in region i that scale the utility derived by households in that region. Amenities can potentially be subject to local congestion externalities as follows:

$$U_i = u_i \cdot L_i^{\beta} \tag{2.20}$$

where u_i are exogenous parameters indicating supply of local amenities (such as parks, lakes, roads, etc.) and $\beta < 0$ measures the strength of the congestion effect.

Such an extension to the base model would leave most equilibrium conditions from (2.2) to (2.9) unchanged. The most important difference will be in the welfare equalization condition (2.10), which now would include additional terms for amenities U_i:

$$\frac{w_i}{P_i} \cdot U_i = \frac{w_j}{P_j} \cdot U_j = W \qquad \text{for all} \quad i, j \in S \tag{2.21}$$

thereby allowing real wages to vary endogenously between locations. The system of equations (2.12) will then also include the amenity terms:

$$\begin{aligned}
\left(w_i \cdot U_i\right)^{1-\sigma} = \left(w_i \cdot u_i \cdot L_i^{\beta}\right)^{1-\sigma} &= W^{1-\sigma} \cdot \sum_s a_s^{\sigma} \cdot \tau_{si}^{1-\sigma} \cdot A_s^{\sigma-1} \cdot w_s^{1-\sigma} \\
w_i^{1-\sigma} \cdot L_i^{\beta(1-\sigma)} &= W^{1-\sigma} \cdot \sum_s a_s^{\sigma} \cdot A_s^{\sigma-1} \cdot \tau_{si}^{1-\sigma} \cdot u_i^{\sigma-1} \cdot w_s^{1-\sigma} \\
W^{\sigma-1} &= \sum_s a_s^{\sigma} \cdot A_s^{\sigma-1} \cdot \tau_{si}^{1-\sigma} \cdot u_i^{\sigma-1} \cdot w_s^{1-\sigma} \cdot w_i^{\sigma-1} \cdot L_i^{\beta(\sigma-1)}
\end{aligned} \tag{2.22}$$

Expenditure shares in (2.14) will then be rewritten as:

$$\pi_{si} = \frac{a_s^\sigma \tau_{si}^{1-\sigma} w_s^{1-\sigma} A_s^{\sigma-1}}{P_i^{1-\sigma}} \;=\; a_s^\sigma \tau_{si}^{1-\sigma} A_s^{\sigma-1} \cdot \frac{w_s^{1-\sigma}}{P_s^{1-\sigma}} \cdot \frac{U_s^{1-\sigma}}{U_s^{1-\sigma}} \frac{P_s^{1-\sigma}}{P_i^{1-\sigma}} \cdot \frac{U_i^{1-\sigma}}{U_i^{1-\sigma}}$$

$$\pi_{si} = W^{1-\sigma} \cdot a_s^\sigma \tau_{si}^{1-\sigma} A_s^{\sigma-1} \cdot \left[\frac{w_s}{w_i}\right]^{1-\sigma} U_i^{\sigma-1}$$

$$\pi_{si} = W^{1-\sigma} \cdot a_s^\sigma \tau_{si}^{1-\sigma} A_s^{\sigma-1} \cdot \left[\frac{w_s}{w_i}\right]^{1-\sigma} u_i^{\sigma-1} L_i^{\beta(\sigma-1)} \tag{2.23}$$

Gross goods shipments will then be given by:

$$y_{si} = \pi_{si} \cdot \frac{w_i}{p_{si}} \;=\; \pi_{si} \cdot \frac{w_i}{\tau_{si} w_s} \cdot A_s$$

$$= \frac{1}{\tau_{si}} \cdot W^{1-\sigma} \cdot a_s^\sigma \tau_{si}^{1-\sigma} A_s^{\sigma-1} \cdot \left[\frac{w_s}{w_i}\right]^{1-\sigma} u_i^{\sigma-1} L_i^{\beta(\sigma-1)} \cdot \frac{w_i}{w_s} \cdot A_s$$

$$\tau_{si} y_{si} = W^{1-\sigma} \cdot a_s^\sigma \tau_{si}^{1-\sigma} \cdot \left[\frac{w_i}{w_s}\right]^{\sigma} u_i^{\sigma-1} L_i^{\beta(\sigma-1)} \cdot A_s^\sigma \tag{2.24}$$

Likewise, the market clearing condition (2.7) will now be a function of amenities as well:

$$A_s L_s = \sum_i \tau_{si} \cdot y_{si} \cdot L_i = W^{1-\sigma} \cdot \sum_i a_s^\sigma \tau_{si}^{1-\sigma} \cdot \left[\frac{w_i}{w_s}\right]^{\sigma} u_i^{\sigma-1} L_i^{\beta(\sigma-1)} \cdot A_s^\sigma \cdot L_i$$

$$w_s^\sigma \cdot L_s = W^{1-\sigma} \cdot \sum_i a_s^\sigma \cdot A_s^{\sigma-1} \cdot \tau_{si}^{1-\sigma} \cdot u_i^{\sigma-1} \cdot w_i^\sigma \cdot L_i^{1+\beta(\sigma-1)}$$

$$W^{\sigma-1} = \sum_i a_s^\sigma \cdot A_s^{\sigma-1} \cdot \tau_{si}^{1-\sigma} \cdot u_i^{\sigma-1} \cdot w_i^\sigma \cdot L_i^{1+\beta(\sigma-1)} \cdot w_s^{-\sigma} \cdot L_s^{-1}$$

$$W^{\sigma-1} = \sum_s a_i^\sigma \cdot A_i^{\sigma-1} \cdot \tau_{is}^{1-\sigma} \cdot u_s^{\sigma-1} \cdot w_s^\sigma \cdot L_s^{1+\beta(\sigma-1)} \cdot w_i^{-\sigma} \cdot L_i^{-1} \tag{2.25}$$

Equations (2.25) and (2.22) together imply that for all $i \in S$, the following must hold:

$$\sum_s \tau_{si}^{1-\sigma} \cdot \frac{a_s^\sigma \cdot A_s^{\sigma-1} \cdot w_s^{1-\sigma}}{a_i^\sigma \cdot A_i^{\sigma-1} \cdot w_i^{1-\sigma}} = \sum_s \tau_{is}^{1-\sigma} \cdot \frac{u_s^{\sigma-1} \cdot w_s^\sigma \cdot L_s^{1+\beta(\sigma-1)}}{u_i^{\sigma-1} \cdot w_i^\sigma \cdot L_i^{1+\beta(\sigma-1)}} \tag{2.26}$$

If the transportation costs are assumed to be symmetric ($\tau_{is} = \tau_{si}$), then equation (2.26) implies the following equality for all pairs $s, i \in S$:

$$\begin{aligned}\frac{a_s^\sigma \cdot A_s^{\sigma-1} \cdot u_s^{1-\sigma}}{a_i^\sigma \cdot A_i^{\sigma-1} \cdot u_i^{1-\sigma}} = \frac{w_s^{2\sigma-1} \cdot L_s^{1+\beta(\sigma-1)}}{w_i^{2\sigma-1} \cdot L_i^{1+\beta(\sigma-1)}} \qquad \Leftrightarrow \\ \left[\frac{w_s}{w_i}\right]^{2\sigma-1} = \left[\frac{a_s}{a_i}\right]^{\sigma} \cdot \left[\frac{A_s}{A_i}\right]^{\sigma-1} \cdot \left[\frac{u_s}{u_i}\right]^{1-\sigma} \cdot \left[\frac{L_s}{L_i}\right]^{-1-\beta(\sigma-1)}\end{aligned} \tag{2.27}$$

Plugging (2.27) in (2.22), we will obtain a non-linear system in labor allocations and parameters:

$$\begin{aligned}W^{\sigma-1} &= \sum_s a_s^\sigma \cdot A_s^{\sigma-1} \cdot \tau_{si}^{1-\sigma} \cdot u_i^{\sigma-1} \cdot L_i^{\beta(\sigma-1)} \\ &\quad \cdot \left[\frac{a_i}{a_s}\right]^{\sigma\hat{\sigma}} \cdot \left[\frac{A_i}{A_s}\right]^{(\sigma-1)\hat{\sigma}} \cdot \left[\frac{u_i}{u_s}\right]^{(1-\sigma)\hat{\sigma}} \cdot \left[\frac{L_i}{L_s}\right]^{-(1+\beta(\sigma-1))\hat{\sigma}} \\ L_i^{\hat{\sigma}-\beta(\sigma-1)(1-\hat{\sigma})} &= W^{1-\sigma} \cdot a_i^{\sigma\hat{\sigma}} \cdot A_i^{(\sigma-1)\hat{\sigma}} \cdot u_i^{(\sigma-1)(1-\hat{\sigma})} \\ &\quad \cdot \sum_s a_s^{\sigma(1-\hat{\sigma})} \cdot A_s^{(\sigma-1)(1-\hat{\sigma})} \cdot \tau_{si}^{1-\sigma} \cdot u_s^{(\sigma-1)\hat{\sigma}} \cdot L_s^{(1+\beta(\sigma-1))\hat{\sigma}}\end{aligned}$$

$$L_i^{\hat{\sigma}(1-\beta\sigma)} = W^{1-\sigma} \cdot a_i^{\sigma\hat{\sigma}} \cdot A_i^{(\sigma-1)\hat{\sigma}} \cdot u_i^{\sigma\hat{\sigma}} \cdot \sum_s \tau_{si}^{1-\sigma} \cdot a_s^{\sigma(1-\hat{\sigma})} \cdot A_s^{\sigma\hat{\sigma}} \cdot u_s^{(\sigma-1)\hat{\sigma}} \cdot L_s^{\hat{\sigma}(1+\beta(\sigma-1))} \tag{2.28}$$

where $\hat{\sigma} = (\sigma-1)/(2\sigma-1)$. While equilibrium labor allocations can technically be solved directly from (2.28), Allen and Arkolakis (2014) show that the equilibrium can be computed as the limit of a sequence defined by

$$f_{i,n+1} = \sum_s K_{si} \cdot f_{i,n}^{\frac{1+\beta(\sigma-1)}{1-\beta\sigma}}$$

which is easier to implement computationally. Here,

$$\begin{aligned}f_i &= L_i^{\hat{\sigma}(1-\beta\sigma)} \qquad \text{and} \\ K_{si} &= \tau_{si}^{1-\sigma} \cdot a_s^{\sigma(1-\hat{\sigma})} \cdot A_s^{\sigma\hat{\sigma}} \cdot u_s^{(\sigma-1)\hat{\sigma}} \cdot a_i^{\sigma\hat{\sigma}} \cdot A_i^{(\sigma-1)\hat{\sigma}} \cdot u_i^{\sigma\hat{\sigma}}\end{aligned}$$

Once labor allocations are solved for and scaled to $\bar{L}$, wages can be computed using

(2.27).

Imputation of local productivities and amenities from wage and employment data

Given data on employment and wages, the extended model, likewise to the base model, allows imputation of local amenities and productivities. Note from (2.27) that we can express productivities in terms of wages, labor, and amenities:

$$\frac{a_i^\sigma A_i^{\sigma-1}}{a_j^\sigma A_j^{\sigma-1}} = \frac{w_i^{2\sigma-1}}{w_j^{2\sigma-1}} \cdot \frac{L_i}{L_j} \cdot \frac{U_i^{\sigma-1}}{U_j^{\sigma-1}} \tag{2.29}$$

We can also rewrite (2.25) as follows:

$$\frac{a_i^\sigma A_i^{\sigma-1}}{a_j^\sigma A_j^{\sigma-1}} = \frac{w_i^\sigma L_i}{w_j^\sigma L_j} \cdot \frac{\sum_s \tau_{js}^{1-\sigma} \cdot w_s^\sigma \cdot L_s \cdot U_s^{\sigma-1}}{\sum_s \tau_{is}^{1-\sigma} \cdot w_s^\sigma \cdot L_s \cdot U_s^{\sigma-1}} \tag{2.30}$$

Combining (2.29) and (2.30), composite amenities can be expressed as a nonlinear system in wages and employment data:

$$\frac{U_i^{1-\sigma}}{U_j^{1-\sigma}} = \frac{\sum_s \tau_{is}^{1-\sigma} \cdot w_i^{\sigma-1} \cdot w_s^\sigma \cdot L_s \cdot U_s^{\sigma-1}}{\sum_s \tau_{js}^{1-\sigma} \cdot w_j^{\sigma-1} \cdot w_s^\sigma \cdot L_s \cdot U_s^{\sigma-1}} \tag{2.31}$$

Relative productivities can then be imputed, using (2.29). While (2.31) can technically be solved directly, the procedure suggested by Allen and Arkolakis (2014) involves much less computational cost, particularly when the number of regions is large. As they have shown, (2.31) implies the following:

$$U_i^{1-\sigma} = \phi \cdot \sum_s \tau_{si}^{1-\sigma} \cdot w_i^{\sigma-1} \cdot w_s^\sigma \cdot L_s \cdot \left[U_s^{1-\sigma}\right]^{-1} \tag{2.32}$$

for an arbitrary scalar ϕ, and the solution can be computed as the convergence of the sequence:

$$
\begin{aligned}
f_{i,n+1} &= \sum_s K_{si} \cdot f_{i,n}^{-1} \qquad\qquad \text{where} \\
f_i &= U_i^{1-\sigma} \qquad \text{and} \\
K_{si} &= \tau_{si}^{1-\sigma} \cdot w_i^{\sigma-1} \cdot w_s^{\sigma} \cdot L_s
\end{aligned}
$$

Spatial impulse responses

Consider again the simple example of an economy on the line as before, which starts out with identical regions that have identical productivities, utility weights, and amenities (i.e. $A_i = a_i = u_i = 1$ for all $i \in S$), as illustrated below. Qualitatively, all the desirable features from the base model carry through, e.g. employment and incomes increase in the booming region while the remaining regions show impulse responses monotonically decreasing with distance away from the epicenter.

Quantitatively, predicted responses of the model are much more in line with the observed data. For instance, booming regions now have increasing nominal wages. Employment responses are likewise much more muted relative to the base case, with 1 percent increase in the relative productivity now causing about 0.8 percent increase in employment of the region in the center.

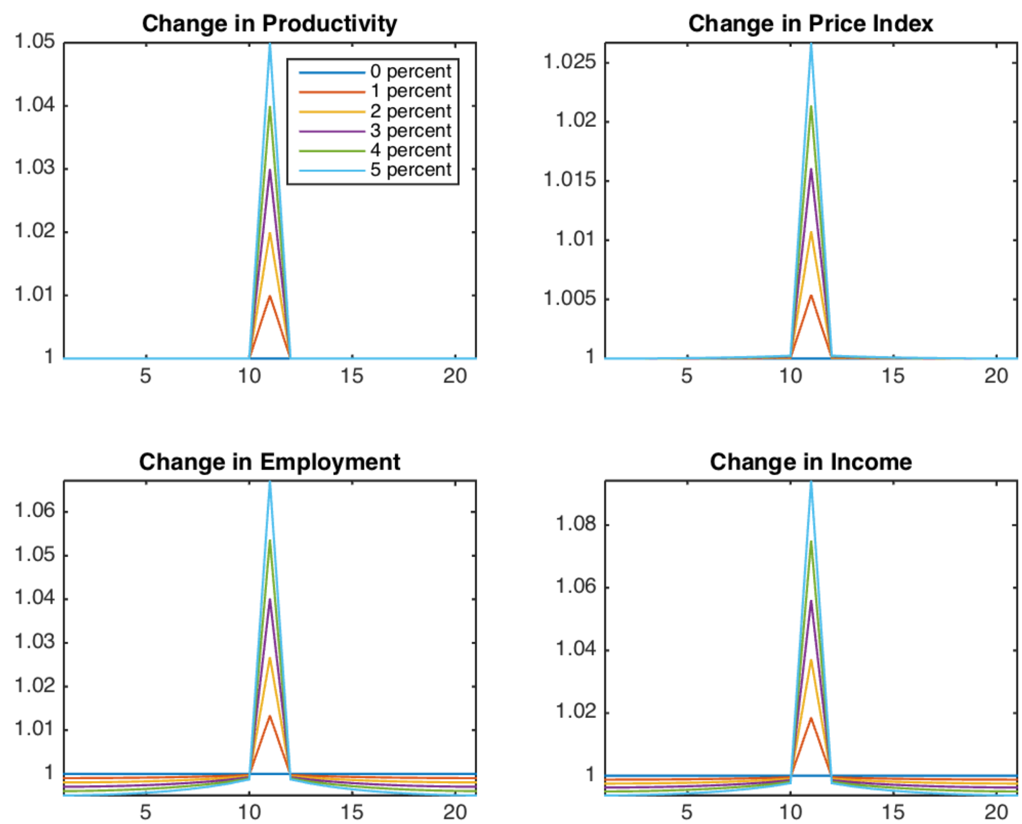


Figure 2.4: Symmetric economy on the line ($a_i = u_i = 1$ for all i, $\sigma = 5$, $\beta = -0.4$)

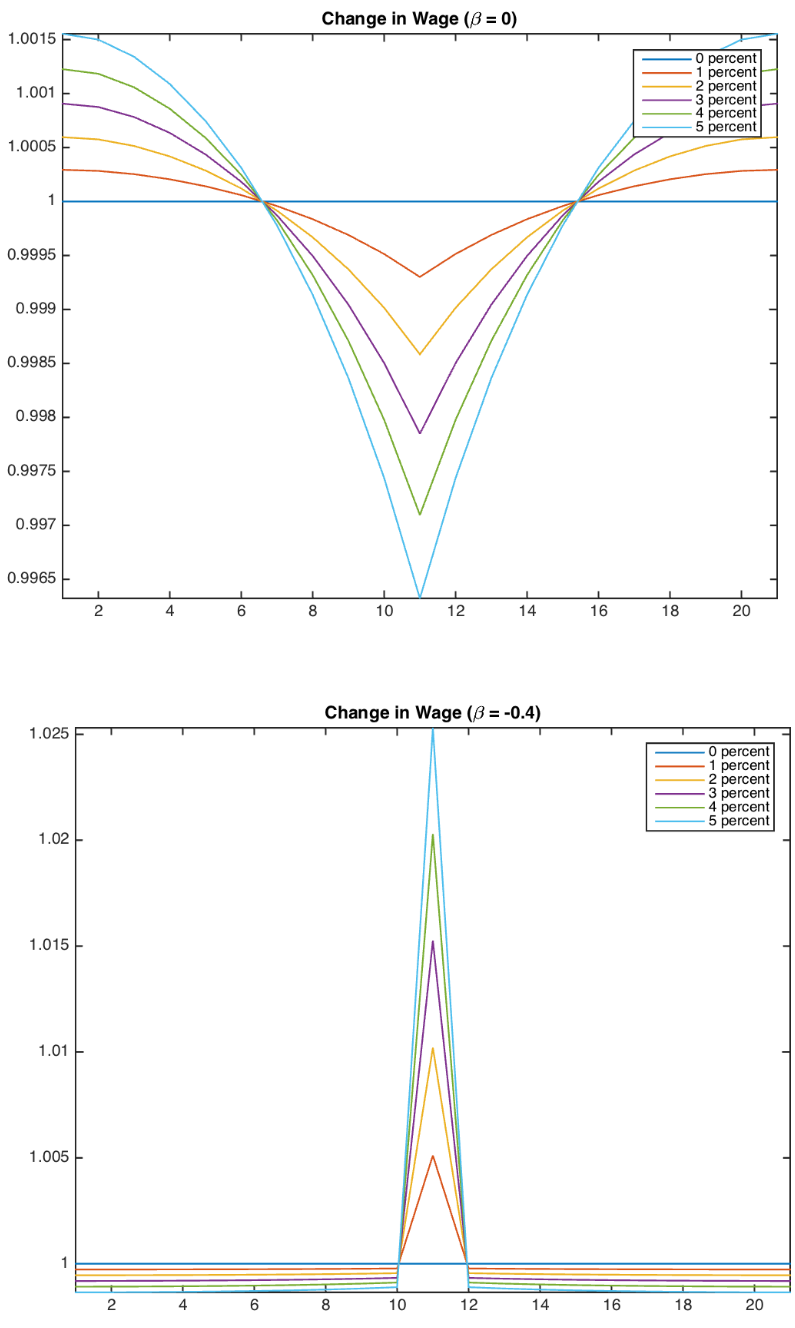


Figure 2.5: Spatial impulse responses with and without congestion effects

Chapter 3

Bakken Shale and Spatial Spillovers

In this chapter, the spatial equilibrium model outlined in Chapter 2 will be applied to the case of the ongoing shale energy boom centered around 12 oil-producing counties in western North Dakota and eastern Montana[1]. The chapter opens with a thesis that the current boom is primarily underpinned by increases in productivity driven by technological innovations in the oil and gas industry. I present stylized facts about the economic performance of this region and report evidence of a spatial pattern of impact. The chapter is concluded with structurally derived estimates of the economic impact.

3.1 Shale Energy Revolution

Technological Revolution

Over the past decade, a technological innovation in the oil and gas industry led to an economic boom in many parts of the United States, where advancements in seismic imaging, horizontal drilling and hydraulic fracturing (fracking) techniques made extraction of previously inaccessible shale oil and gas resources economically viable.

[1] Richland, Roosevelt, and Sheridan Counties in Montana and Billings, Burke, Divide, Dunn, Golden Valley, McKenzie, Mountrail, Stark, and Williams Counties in North Dakota.

One such region strongly affected by the shale energy boom is the Bakken area straddling western North Dakota and eastern Montana, named after the shale formation underground. Since 2003, when shale oil activity first started, the area experienced extremely rapid growth in employment and wages and resulted in record low unemployment rates. Within a decade, oil production in the area surged to pass 1 million barrels per day, currently comprising over 13 percent of total U.S. production.

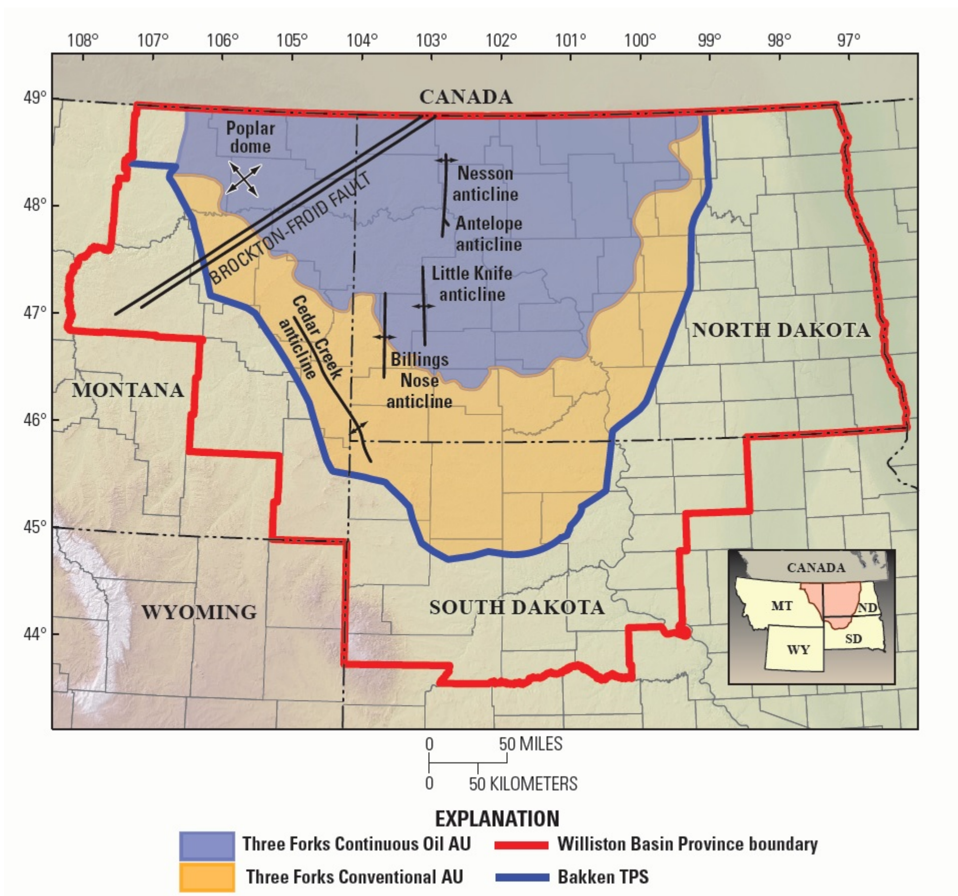


Figure 3.1: Bakken and Three Forks Formations, U.S. Geological Survey

The first development that contributed in a significant way to the shale energy revolution was advancements in the three-dimensional seismic imaging technologies, aided by the advent of high performance computing. In an overview of these developments, Cartwright and Huuse (2005) note how the transition from 2D to 3D seismic surveying dramatically improved the resolution, reducing a typical grid spacing from a kilometer down to 25 meters or less (or from about 9 football field lengths down to 27 yards).

These technological improvements led to a sharp increase in the success rate of crude oil and gas well drilling. According to the Energy Information Administration data, the 'dry hole' percentage of drilled wells dropped significantly since 1997, when incidentally many of the seminal papers on 3D seismic imaging seem to have been published. The percentage of exploratory wells drilled that failed to yield significant amounts oil or gas nearly halved from about 70 percent in 1997 down to about 35 percent in late 2000s.

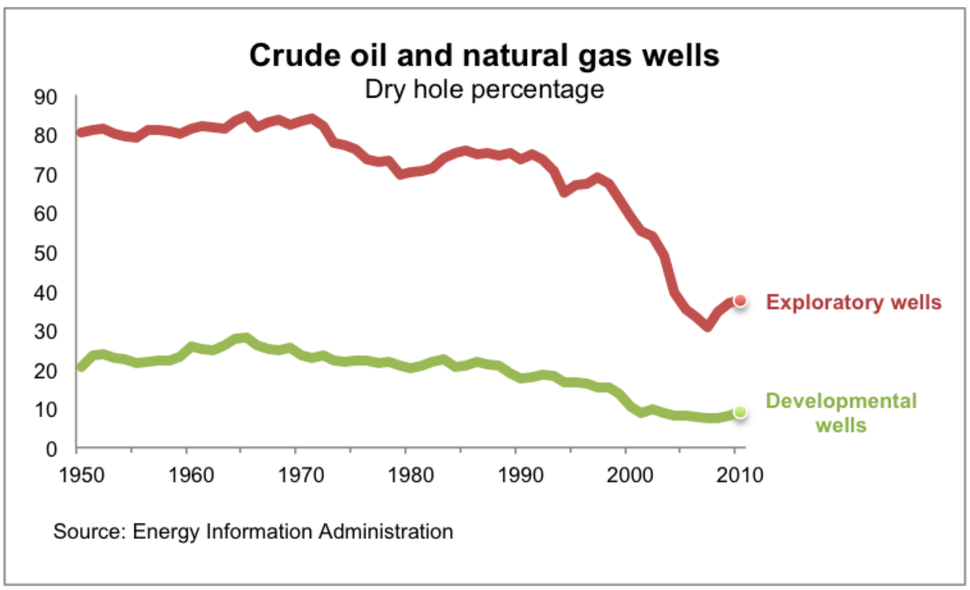


Figure 3.2: Failure rate of oil and gas well drilling significantly decreased

Another important contributing factor to the shale energy revolution was advancements in the horizontal drilling technologies. While directional drilling technology had been around for decades, Downton and Hendricks (1999) survey how development of new rotary steerable tools, first commercially introduced in 1996, have increased productivity and reduced costs of drilling directional and horizontal wells. These developments depended on accurate surveying systems as well as improvements in downhole devices, allowing real-time communication with the surface as well as fine precision of steering control.

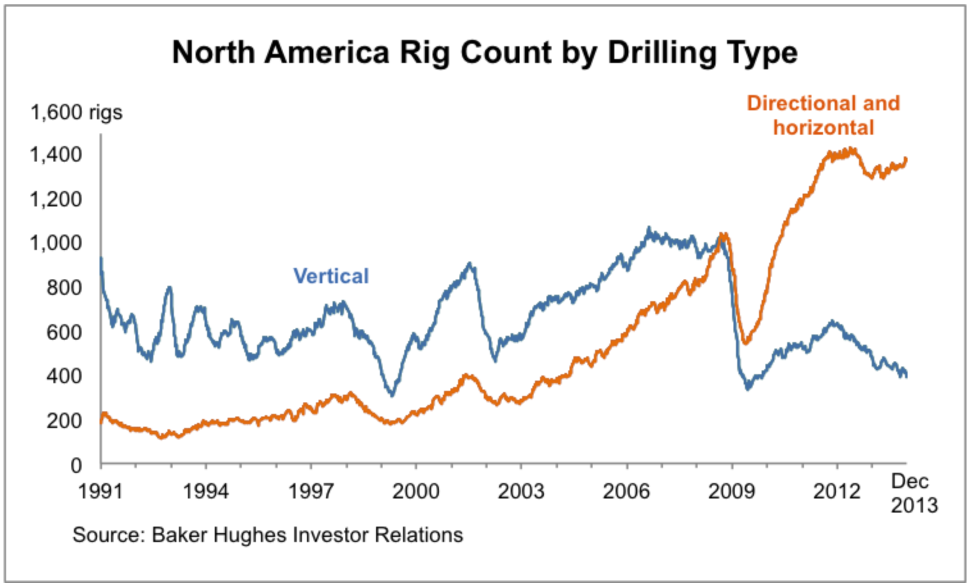


Figure 3.3: Growth in horizontal wells far outpaced growth in conventional wells

Greater precision of well trajectories increased rate of penetration (speed of drilling), lowering the overall costs per well. Downton and Hendricks (1999), for example, note how the usage of a rotary steerable system in Norway's Njord field cost $1 million less than the previous well drilled using conventional methods in the same field because it cut well construction time by half. Furthermore, rotary steerable systems meant that each well can now reach more targets, further increasing productivity of each well drilled. As a result, directional and horizontal type drilling rigs grew from occupying a relatively small share of the U.S. total in the 1990s to dominating the industry by the late 2000s, according to Baker Hughes data.

Lower risk of collision (drilling into another wellbore) has been another important benefit of this new technology, which allowed for "pad" drilling whereby multiple wells are drilled from the same pad. This resulted in lower infrastructure costs, greater efficiencies in terms of reduced rig downtime as well as a smaller footprint on the surface. Energy

Information Administration (2012), for instance, notes how operators in the Eagle Ford shale formation reduced average times of drilling a horizontal well from 23 days in 2011 down to 19 days in 2012, thanks to pad drilling and moveable rigs.

Finally, the third piece underpinning the shale energy boom was improvements in hydraulic fracturing methods of stimulating well production, which generally involve injecting fluids into the target shale formation at high pressures to fracture the porous rock in order to release oil and natural gas trapped within. Although hydraulic fracturing methods have been in use for decades prior to the current boom, Fitzgerald (2013) notes that the first success after a long history of experimentation of fracking was recorded in 1998 in Barnett shale of Texas and lists four technical innovations that distinguish contemporary fracking from its predecessors. The first difference is substantially larger volumes of fluid and proppants used, which can involve injecting millions of gallons of water as well as millions of pounds of proppants such as sand into each well. The second is that contemporary fracking combines water with gelling agents ("slickwater" fracking) to enhance flow of large quantities of proppants and improve permeability. The third is that contemporary fracking is performed in multiple stages, thanks to the ability to isolate sections of the wellbore for each individual fracking stage, thereby providing greater control over the process and more power to each stage. Lastly, significant improvements went into the optimal mix of additives to maximize well production, which differ by company and characteristics of the target formation.

As a result, "tight" or unconventional oil production in the U.S. rapidly expanded over the past decade, more than compensating for the falling production from conventional oil wells. According to the Energy Information Administration data, tight oil production grew from about 0.4 million barrels per day in January 2000 to 4.6 million barrels per day in February 2015, and its share of total output rose from about 6 percent in 2000 to about 45 percent in 2014. The Bakken region's oil production share rose from under 1 percent of U.S. total in 2000 to over 13 percent by the end of 2014.

There are also indications that drilling methods and oil and gas extraction techniques

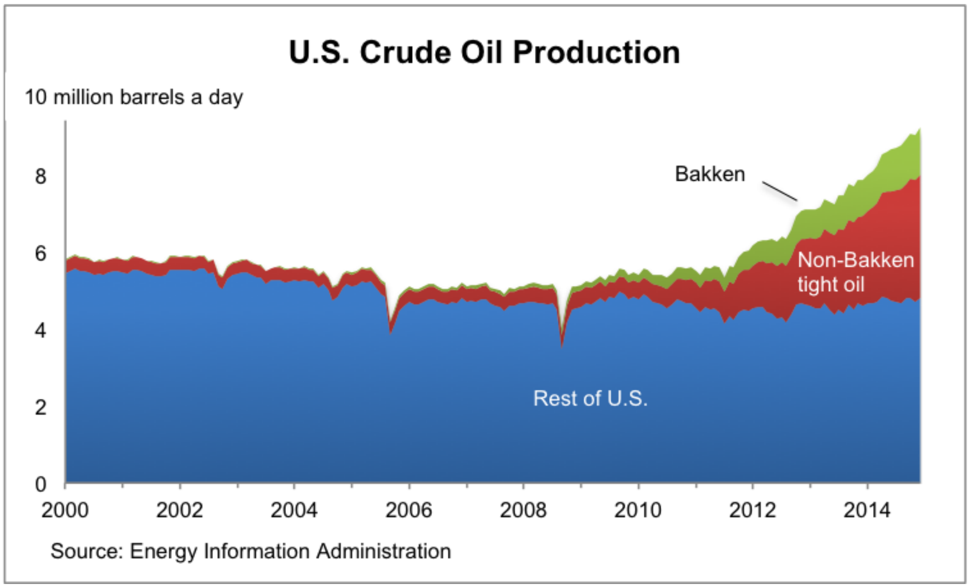


Figure 3.4: Shale oil reversing decline in oil production from conventional wells

continually improved over the past decade, leading to a sustained growth in productivity. For instance, well-level production data published by North Dakota's Industrial Commission in its Monthly Production Reports show progressively rising peak oil production levels of wells drilled in years since 2003. Although these wells exhibit steeper decline curves with oil output typically falling sharply within the first 12 months of production, peak production of an average well rose from about 2,000 barrels of oil per month in 2004 to over 10,000 barrels in 2014.

While the exact source of these more recent productivity increases is difficult to identify precisely, Cochener (2010) offers a number of explanations for increased drilling efficiency: (i) reduced non-productive time of drilling rigs thanks to use of faster to assemble rigs and more efficient work practices; (ii) increased rate of penetration (speed of drilling) from improved drill bits, use of synthetic fluids in drilling muds, and utilizing smaller bore holes; (iii) "mixed fleet" management strategy which involves sequential use of different types of rigs specialized in different stages of drilling a well, in a fashion resembling an assembly line; and (iv) better overall management and planning.

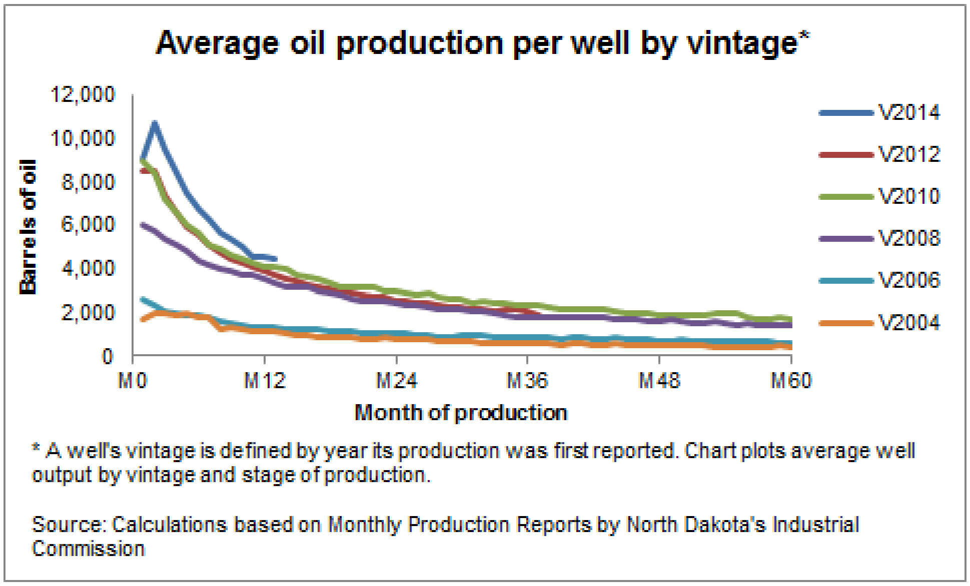


Figure 3.5: Productivity of newly drilled Bakken wells has been continuously rising over the past decade

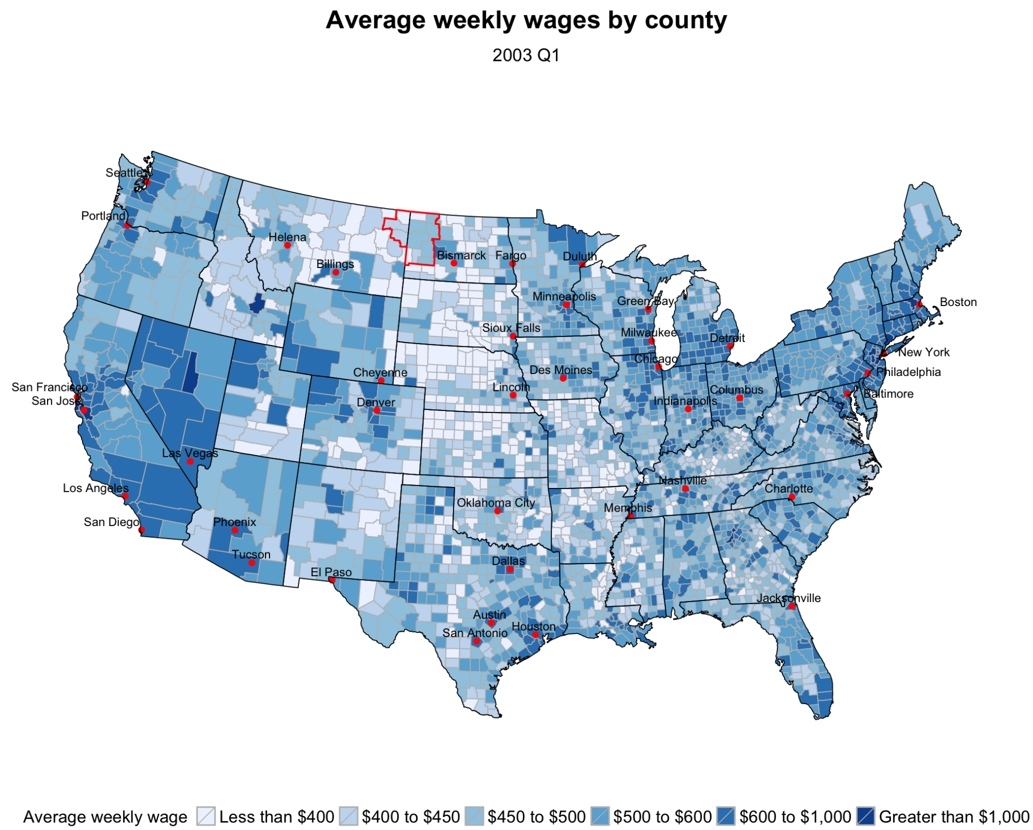
Learning-by-doing is another possible explanation as firms gain experience from experimenting with different lengths of horizontal shafts, numbers of fracking stages, and different mixes of fracking fluids, all tailored to each particular location. Lastly, productivity growth was partly driven by continued innovations in the industry, such as the technique of drilling two or more parallel wells and perforating at alternate intervals, which allows for a high-density network of fractures between wells; use of ultraviolet light to kill bacteria that impede production; wellsite recycling of wastewater which allows reuse of water; and safer and better fracking chemicals.

Whatever the source of the underlying productivity growth, it has clearly been the driving force behind the ongoing shale energy boom that has revolutionized the industry, changed the U.S. energy landscape, and transformed entire regions.

Spatial Pattern of Impact

While the economic impact of the oil boom in the immediate Bakken area is most noticeable, much less is generally known about its spillover effects on neighboring regions. However, there are some indications that the economic effect from the shale energy boom on surrounding areas may have a spatial pattern with counties closer to the Bakken experiencing stronger impact relative to those farther out.

Prior to the oil boom, Bakken counties used to have some of the lowest wages in the nation. According to the Quarterly Census of Employment and Wages (QCEW) data by the Bureau of Labor Statistics, weekly wages in most Bakken counties averaged well below \$500 dollars at the beginning of 2003. A decade later, however, the situation reversed with the Bakken counties now having some of the highest average wages in the nation. As of the second quarter of 2014, wages in most Bakken counties averaged well over \$1,000 dollars per week.



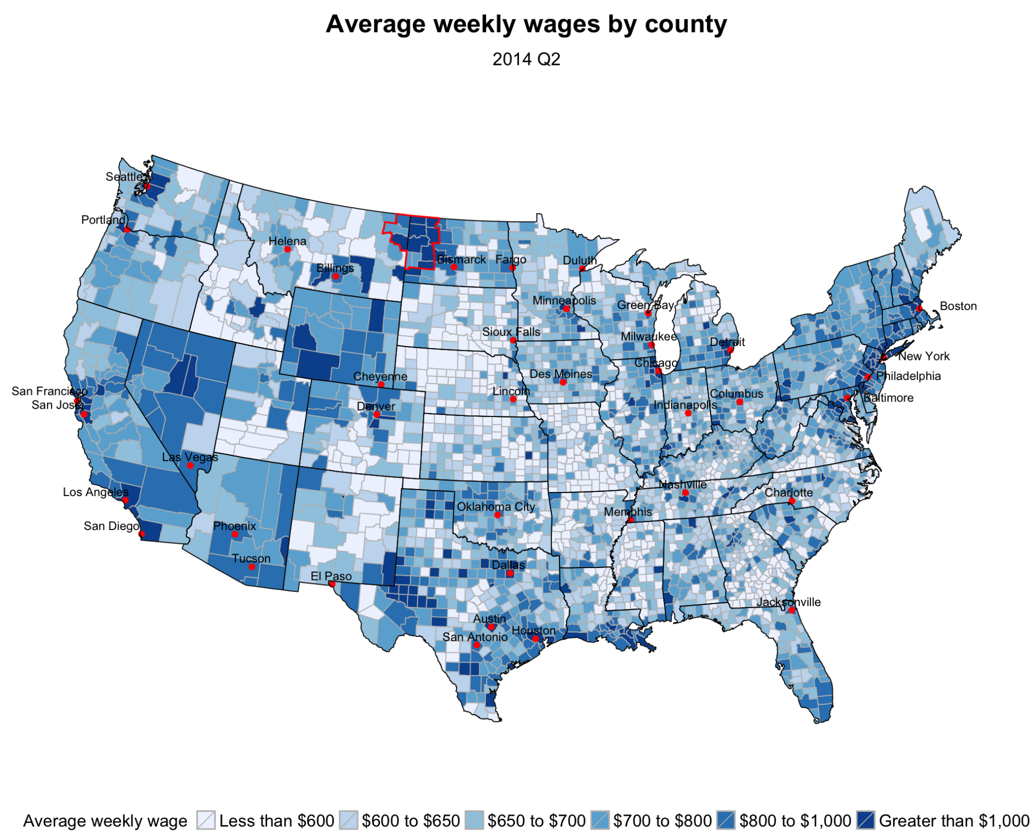


Figure 3.6: Based on Bureau of Labor Statistics (QCEW)

Bakken counties are among the very few in the nation that had their average weekly wages more than double since 2003. Moreover, the heat map of U.S counties also suggests strong wage growth in surrounding regions as well, with counties closer in distance to the Bakken clearly experiencing higher wage growth relative to those farther out over this period.

Notably, most of the counties that are shown to have experienced as strong a wage growth as Bakken counties during this period are in or around other shale oil and gas producing areas such as those in the Eagle Ford and Permian Basin shale plays in Texas.

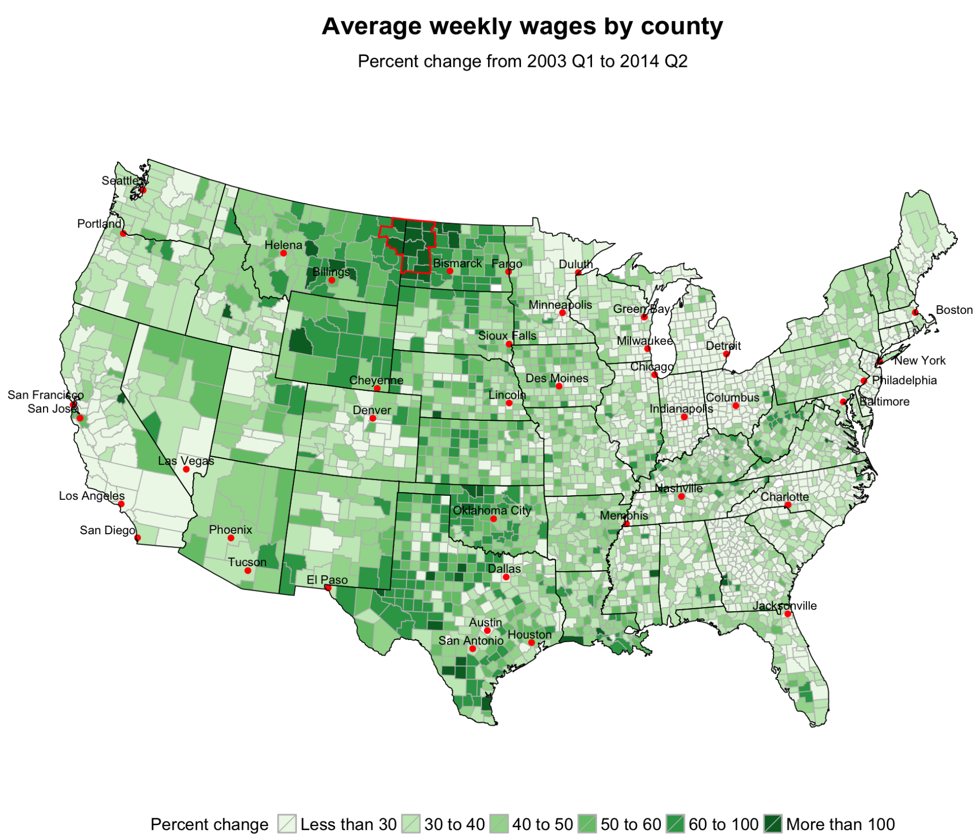
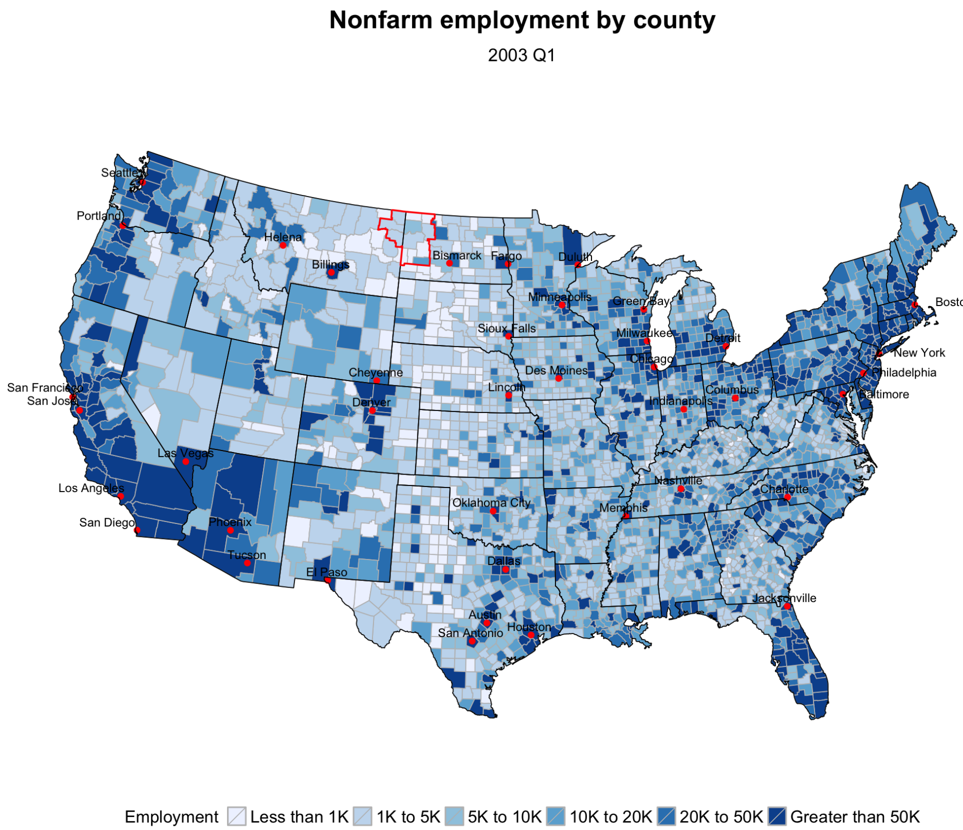


Figure 3.7: Wage growth strongest in the Bakken and dissipating with distance from the epicenter



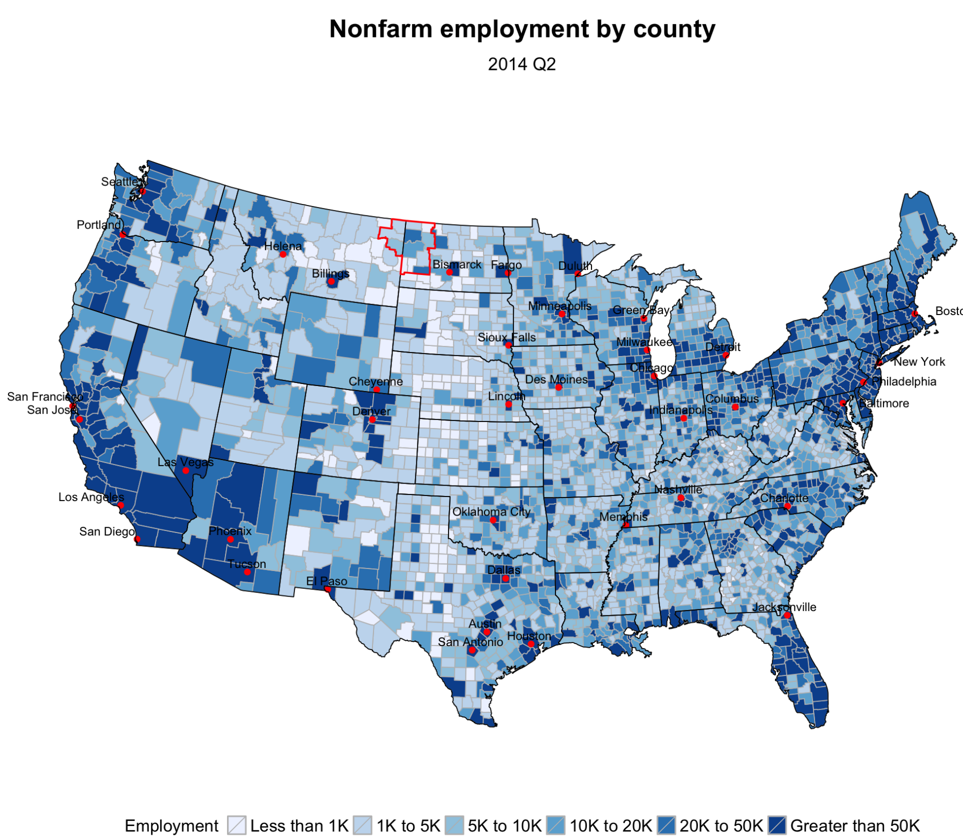


Figure 3.8: Based on Bureau of Labor Statistics (QCEW)

In 2003, Bakken was a sparsely populated and mostly rural region. North Dakota's Stark county was the largest with just over 10,000 workers, and employment in the entire region totaled about 36,000 workers. As of 2014 Q2, however, Bakken employment more than doubled to cross just above the 100,000 mark. Employment in the Williams county at the core of the Bakken area more than quadrupled, overtaking Stark to become the largest among the twelve.

More importantly, employment growth over this period seems to exhibit broadly similar patterns to the wage growth. Just as with wage growth, counties in the immediate vicinity of the Bakken generally seem to have also experienced higher than average increases in employment since the start of the oil boom.

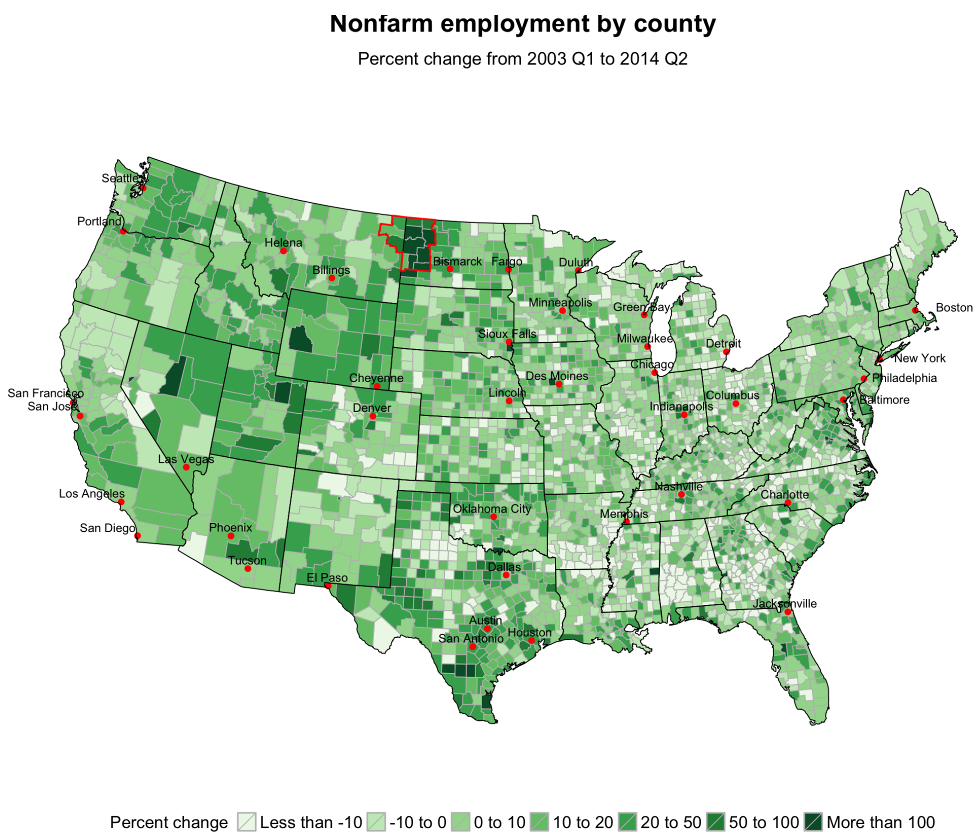


Figure 3.9: Employment growth concentrated in oil producing counties

A standard way of measuring spatial correlation in the data is to compute the Moran's I statistic, which relates the degree to which a county's characteristics such as wage growth is correlated with the weighted average of its neighbors. The data for neighboring counties is typically weighted using a spatial weight matrix W, which can be based on measures of contiguity or some metric of distances between them.

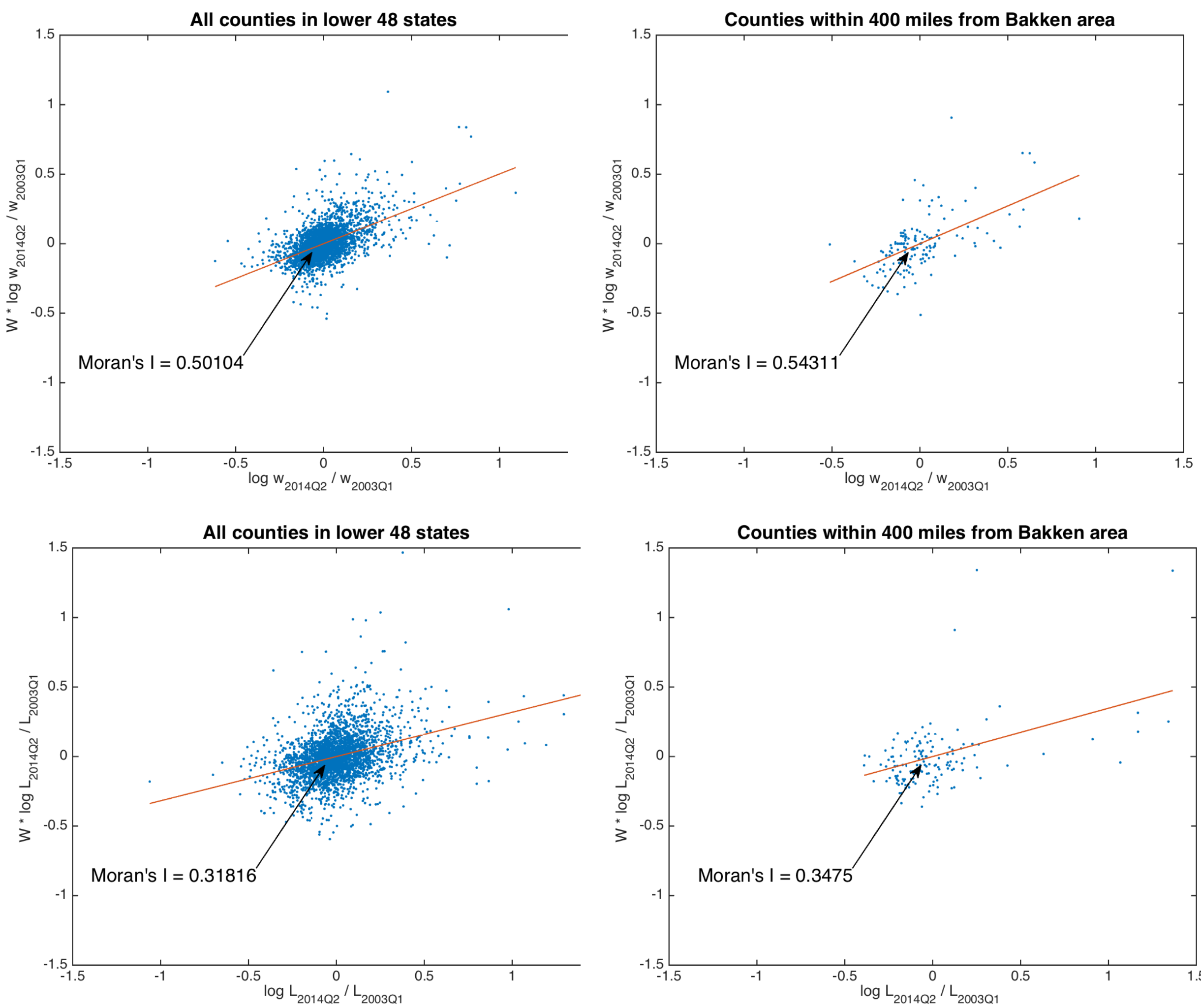


Figure 3.10: Employment and wage growth spatially correlated for counties surrounding the Bakken

Moran scatter plots in Figure 3.10, for instance, chart wage and employment growth (deviations from the mean) data by county for the period from 2003Q1 to 2014Q2 against a spatially-weighted average statistic for the rest of counties. The plots show clear signs of spatial correlation, which is stronger for wage growth relative to employment growth data. Moreover, Moran's measure of spatial correlation is positive, indicating spatial clustering of counties in terms of their economic performance, and the measure gets higher if the sample is restricted to counties within 400 miles from the Bakken area.

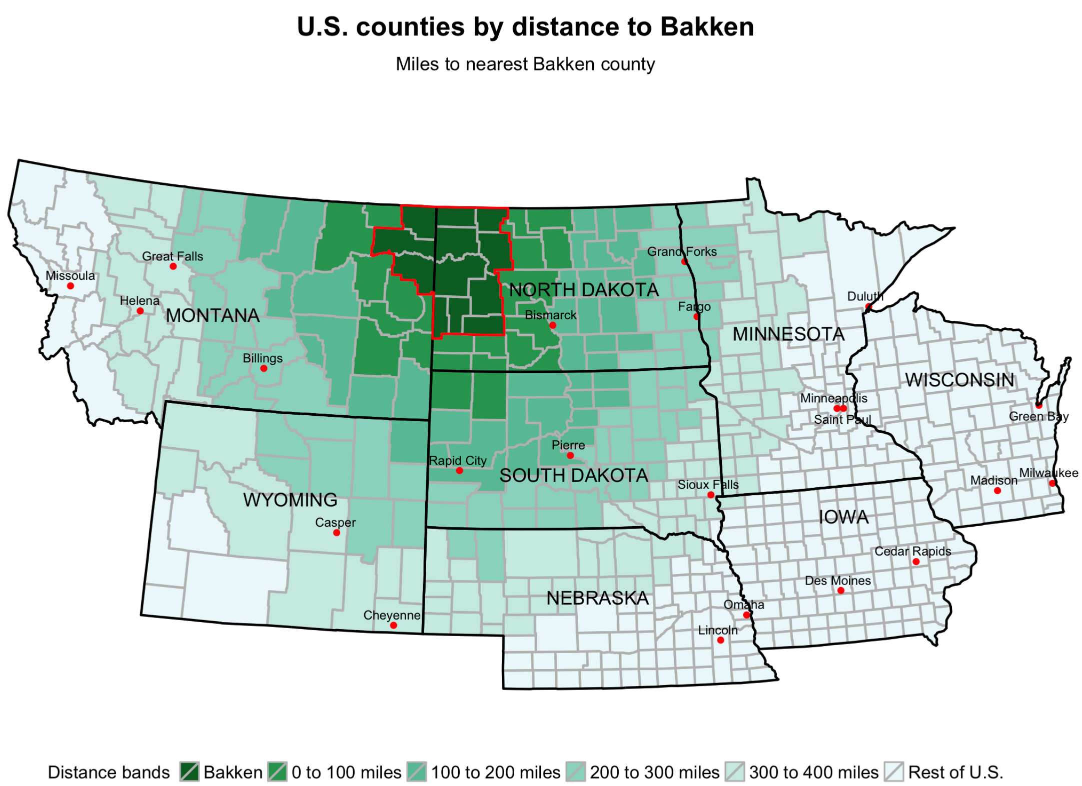


Figure 3.11: 100-mile band rings around the Bakken

A more convenient way of visualizing this pattern, to be repeatedly used in this chapter, would be to follow Batbold and Grunewald (2013), drawing concentric circles in 100-mile increments around the Bakken area and plotting average measures of economic performance for counties in each band. Accordingly, these concentric bands would partition U.S. counties into 6 groupings, as outlined in Figure 3.11 above, for each of which we can plot the average wage and employment statistics.

Not surprisingly, core Bakken counties as a group experienced highest growth rates in employment and wages, approximately tripling in each case (Figure 3.12). More tellingly, counties in the next band (0 to 100 mile distance away from the Bakken) show the second highest growth in both employment and wages, with the effect generally growing weaker with the distance away from the epicenter. All four bands had much stronger growth in both employment and wages relative to the rest of the U.S.

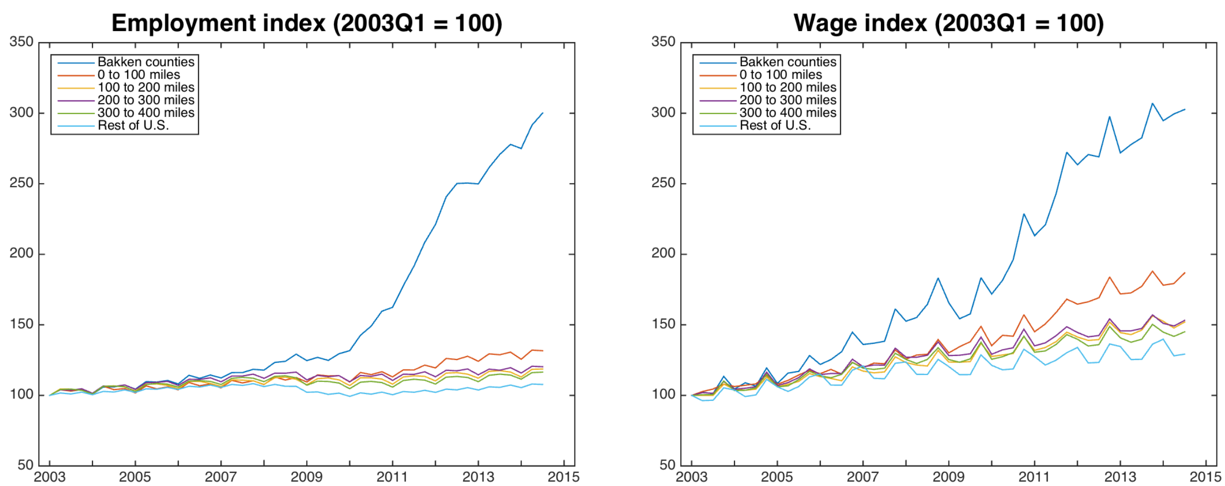


Figure 3.12: Employment and wage growth stronger in counties closer to the Bakken

While such purely descriptive statistics cannot give a definitive answer on the extent of spatial spillovers, they do suggest that any empirical estimation of the impact must take into account the spatial dimension of the problem.

Existing Estimates of the Economic Impact

As discussed in the introduction chapter, input-output analysis and difference-in-differences estimators are most commonly encountered methods of analyzing the economic impact of a mining boom. In the first category, Tunstall et al (2014) estimated a total economic impact of $87 billion in output added as well as 155,000 full-time equivalent jobs supported by the oil production activity in 21 counties directly or indirectly affected by the activity on the Eagle Ford shale in Texas. The core 15 oil-producing counties are estimated to have accounted for 115,000 of these jobs. Brown et al (2013) from Montana Department of Transportation use similar methodology to analyze direct and indirect impacts of the increased oil production activity on industries like oil extraction, construction, and transportation, and arrive at a forecast of 10,000 to 30,000 gain in population as a direct result of the oil activity in eastern Montana.

Marchand (2013) uses a variation of the DID method to estimate the economic impact of the oil boom on the Canadian side of the border and concludes that energy booms lowered poverty rates and increased inequality in Western Canada. Brandt (2013) also uses DID to estimate a 9 percent employment increase in Bakken counties compared to 5 percent increase in non-Bakken counties while finding no difference in wage growth rates between the two county groups post boom.

To the best of my knowledge, there have been no studies to date that analyze the shale oil's impact using a structural model.

3.2 Bakken Impact

Data and Methodology

Methodology

The general procedure for estimating the economic impact is as follows. First, using employment and wage data, implied distribution of relative productivities are backed out. Second, computed productivities are scaled to match the aggregate output. Third, counterfactual paths are computed for productivities in each county in the dataset. Finally, differences between counterfactual employment and wage paths implied by the model are compared to actual data series as a measure of the economic impact.

For the imputation procedure, county-level employment and wage data are obtained for 3,109 counties in the lower 48 U.S. states from the Quarterly Census of Employment and Wages (QCEW) by the Bureau of Labor Statistics. The analysis is done on the quarterly series from first quarter 1990 to second quarter 2014.

Estimation of parameters

The imputation procedure outlined in Chapter 2 requires parameter estimates for $(a_i, \sigma, \tau_{ij}, \beta)$. For the empirical analysis part of this chapter, the transportation costs τ_{ij} are assumed to be proportional to geographical distances between regions.

Geographic distances are approximated using county centerpoint longitude and latitude coordinates from the Census Bureau's TIGER Shapefiles. The Earth's curvature is taken into account using haversine formula, assuming a perfectly spherical shape to the planet. The analysis is focused on counties within the 400 mile radius around the Bakken area, which include counties in Iowa, Minnesota, Montana, Nebraska, North Dakota, South Dakota, and Wyoming, although the model is solved for all counties in lower 48 states.

Using equations (2.5) and (2.6), we can write the ratio of expenditure shares as a function of transportation costs:

$$\frac{\pi_{is}}{\pi_{js}} = \left(\frac{a_i}{a_j}\right)^{\sigma} \left(\frac{\tau_{is}}{\tau_{js}}\right)^{1-\sigma} \left(\frac{w_i}{w_j}\right)^{1-\sigma} \left(\frac{A_i}{A_j}\right)^{\sigma-1}$$

$$\log\left[\frac{\pi_{is}/\pi_{js}}{\pi_{ii}/\pi_{jj}}\right] = (1-\sigma)\log\left[\frac{\tau_{is}}{\tau_{js}}\right] \tag{3.1}$$

Expenditure shares π_{si} can be estimated using Department of Transportation's 2007 Commodity Flow Survey (CFS) data for the lower 48 states in the U.S. If the CFS data were organized into a matrix with origin locations indexed in rows and destination locations in columns, π_{si} would correspond to the share of element in row s in the total sum for column i of the commodity flow matrix, i.e. represent the share of shipments from region s in total shipments delivered to region i.

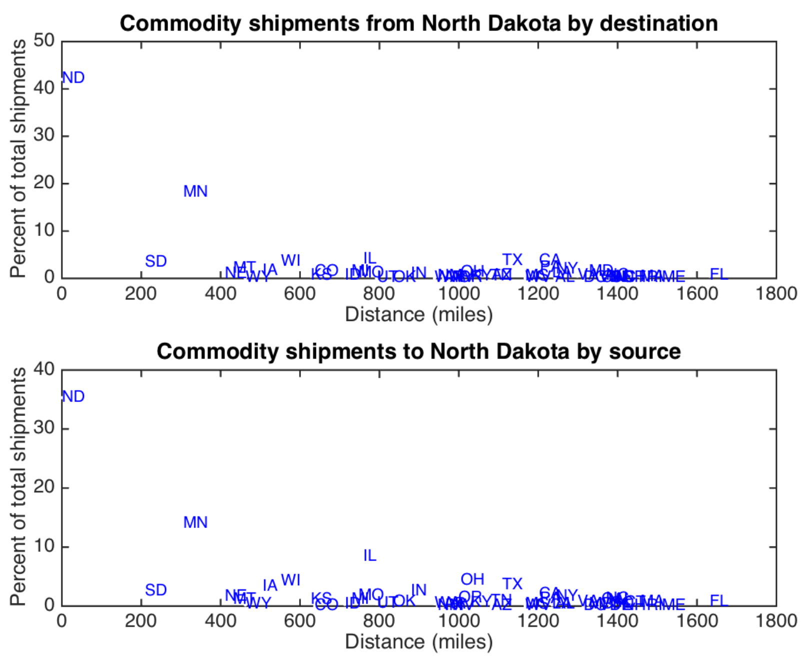


Figure 3.13: Based on the Department of Transportation's 2007 Commodity Flow Survey Data

If we took North Dakota as the destination region, for example, then $\pi_{ND,ND}$ would equal about 0.353, indicating that about 35 percent of total commodity flows shipped to North Dakota came from within the state. Minnesota then would account for the next largest share of commodity flows into North Dakota with $\pi_{MN,ND} = 0.139$ or about 14 percent.

The expenditure shares obtained from the CFS data can be used in equation (3.1) to estimate Armington elasticities, which are obtained under two alternative assumptions about the relationship between transportation costs and geographic distances. If iceberg costs τ are assumed to scale linearly with (normalized) geographical distances, then the Armington elasticity is estimated at $\sigma = 5.3$, whereas if they are assumed to be exponential in distance, then σ is estimated at around 3.9. Intuitively, exponential costs penalize shipments of goods over longer distances, which prompts the estimation procedure to infer lower substitutability between products from different regions. Of the two, the linear transportation cost model appears to have a better fit as indicated by a smaller sum of squared residuals.

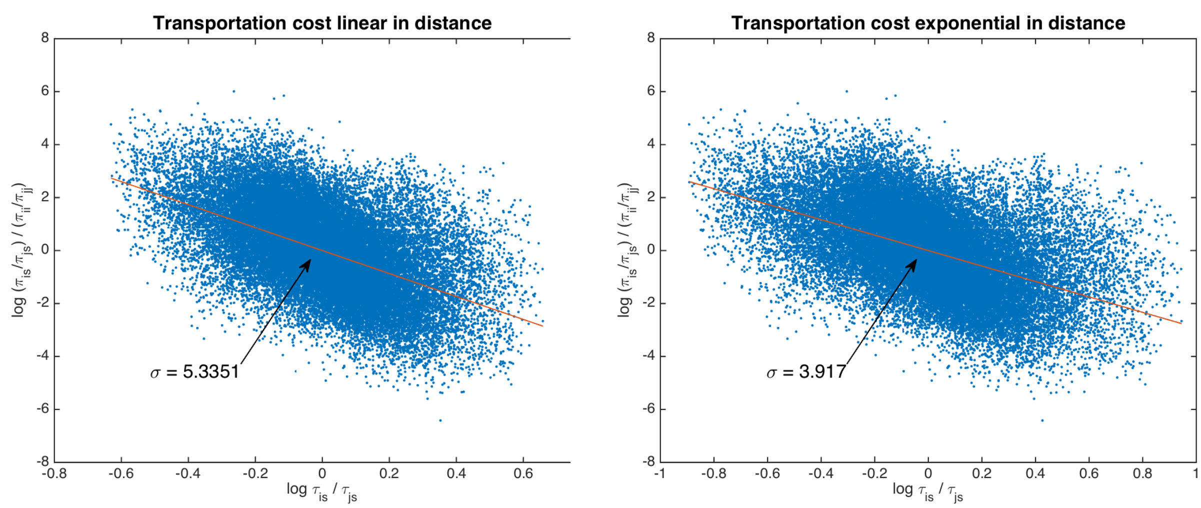


Figure 3.14: Armington elasticity estimates under linear and exponential transportation cost assumptions

For the rest of this chapter, the assumed Armington elasticity will be $\sigma = 5$.

Next, to estimate the exogenous amenity parameters, consider again the welfare equalization condition (2.21):

$$\frac{w_i/P_i}{w_j/P_j} = \frac{U_j}{U_i} \;=\; \frac{u_j}{u_i} \cdot \frac{L_j^\beta}{L_i^\beta}$$

$$\log \frac{w_i/P_i}{w_j/P_j} = -\log \frac{u_i}{u_j} - \beta \log \frac{L_i}{L_j} \tag{3.2}$$

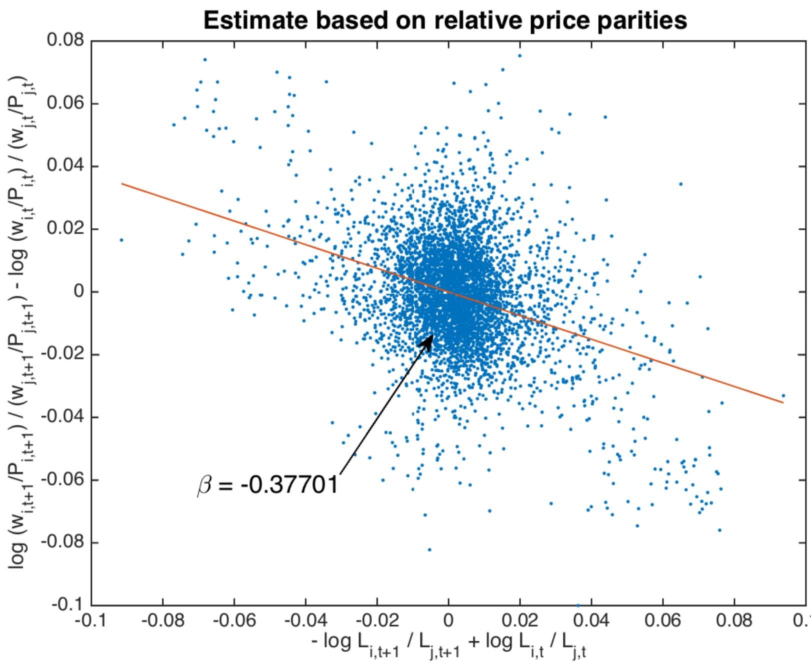


Figure 3.15: Estimate of the congestion parameter β

Congestion parameter β can be estimated by differencing (3.2) as follows:

$$\log \frac{w_{i,t+1}/P_{i,t+1}}{w_{j,t+1}/P_{j,t+1}} - \log \frac{w_{i,t}/P_{i,t}}{w_{j,t}/P_{j,t}} = \beta \left[\log \frac{L_{i,t}}{L_{j,t}} - \log \frac{L_{i,t+1}}{L_{j,t+1}} \right] \tag{3.3}$$

using state-level employment and wage data from the QCEW and relative price parity data from the Bureau of Economic Analysis, which are available at an annual frequency for years 2008 to 2012 and assuming relative amenities u_i/u_j are time-invariant. The resulting estimate from the procedure is $\beta = -0.4$, which is very close to the estimate reported in Allen and Arkolakis (2014).

The above method depends on a strong assumption that relative amenities are time-invariant. An alternative way of estimating β would be to give a specific interpretation to the congestion externality, following Rosen-Roback tradition. Suppose, for instance, that households' preferences are given by:

$$W_i = Y_i^{1-\alpha} \cdot x_i^{\alpha} \tag{3.4}$$

where Y_i is a CES (constant elasticity of substitution) composite of traded goods as in (2.1) and x_i is the household's consumption of local (non-traded) goods such as housing. Given the Cobb-Douglas formulation, expenditure shares will be proportional to α:

$$P_i Y_i = (1-\alpha) \cdot w_i \qquad \text{and} \qquad Q_i x_i = \alpha \cdot w_i \tag{3.5}$$

where Q_i denote region-specific prices of non-traded goods such as housing rental prices. Note that (3.5) implies the following expression for the households' welfare function:

$$W_i = (1-\alpha)^{1-\alpha} \cdot \alpha^{\alpha} \cdot \frac{w_i}{P_i^{1-\alpha} Q_i^{\alpha}} \tag{3.6}$$

In a spatial equilibrium, welfare is equalized across regions so that:

$$W = (1-\alpha)^{1-\alpha} \cdot \alpha^{\alpha} \cdot \frac{w_i}{P_i^{1-\alpha} Q_i^{\alpha}} \quad \text{for all } i \qquad \Leftrightarrow \qquad \log \frac{w_i}{P_i} = c - \alpha \cdot \log \frac{P_i}{Q_i} \tag{3.7}$$

for some constant c. Using Regional Price Parity series from the BEA, one can proxy P_i by the ''RPPs: Goods" series and Q_i by ''RPPs: Services: Rents" series to estimate the following equation implied by (3.7):

$$\log \frac{w_{it}}{P_{it}} = c_t - \alpha \cdot \log \frac{P_{it}}{Q_{it}} \tag{3.8}$$

which can be estimated as a pooled regression model with time-fixed effects.

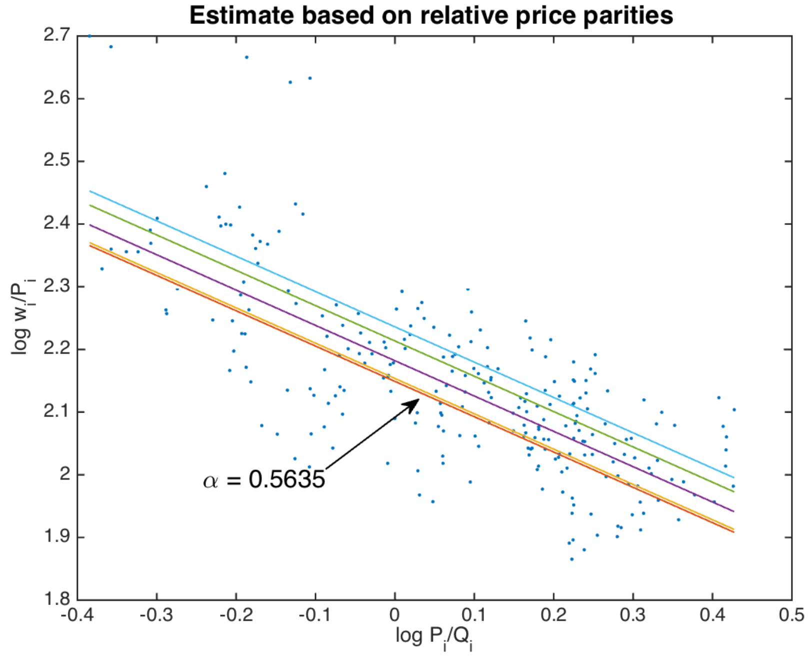


Figure 3.16: Estimated share of household expenditures on non-traded goods

The estimate above implies a representative U.S. household allocates about 56.3 percent of its total expenditures on non-traded goods and services, which is broadly consistent with the share of Personal Consumption Expenditures spent by U.S. households on services (net of financial services, which may be less local). Suppose further that local goods are in fixed supply in any given period so that:

$$x_i L_i = X_i$$

where X_i, the total supply of non-traded goods such as housing, is assumed to be exogenous. Households' optimal decision given in (3.5) will then imply the following price relationship for local goods:

$$Q_i = \alpha \cdot \frac{w_i L_i}{X_i} \tag{3.9}$$

Welfare equalization condition (3.6) then can be rewritten as:

$$W = (1-\alpha)^{1-\alpha} \cdot \alpha^{\alpha} \cdot w_i \cdot P_i^{\alpha-1} \cdot \left[\alpha \cdot \frac{w_i L_i}{X_i}\right]^{-\alpha}$$
$$\hat{W} = \frac{W^{\frac{1}{1-\alpha}}}{(1-\alpha)} = \frac{w_i}{P_i} \cdot \left[X_i^{\frac{\alpha}{1-\alpha}} \cdot L_i^{-\frac{\alpha}{1-\alpha}}\right] \tag{3.10}$$

Note that (3.10) is exactly isomorphic to (2.21) of the extended model where the exogenous amenities correspond to the housing supply $\left(U_i \equiv X_i^{\alpha/(1-\alpha)}\right)$ and the congestion parameter corresponds to an expression of expenditure shares on non-traded goods $\left(\beta = -\alpha/(1-\alpha)\right)$. Given this interpretation, the congestion parameter $\beta = -0.5635/(1-0.5635) = -1.3$, much larger than the previous estimate.

The above estimate represents an upper bound on the range of empirically defensible estimates for the congestion parameter given the specific interpretation of the congestion externality outlined above. If α were interpreted as the share of household income spent on housing and utilities only, then the corresponding share in total Personal Consumption Expenditures of U.S. households averages at about 18.4 percent, according to the BEA data from the 2006-2013 period, which would correspond to a congestion parameter $\beta = -0.23$. Depending on one's assumptions on the degree to which services are tradeable across locations, the congestion parameter can then reasonably range between -0.2 and -1.3.

The only missing piece left to estimate are then the utility weights a_i. Theoretically, utility weights can be computed from (2.5) and (2.6):

$$\frac{\pi_{si}}{\pi_{ii}} = \left(\frac{a_s}{a_i}\right)^{\sigma} \left[\tau_{si} \cdot \frac{w_s/A_s}{w_i/A_i}\right]^{1-\sigma} = \left(\frac{a_s}{a_i}\right)^{\sigma} \left[\tau_{si} \cdot \frac{w_s L_s/A_s L_s}{w_i L_i/A_i L_i}\right]^{1-\sigma} \tag{3.11}$$

Rearranging (3.11), we can back out relative utility weights as a function of previously estimated parameters and empirical counterparts as follows:

$$\frac{a_s}{a_i} = \left(\frac{\pi_{si}}{\pi_{ii}}\right)^{\frac{1}{\sigma}} \left[\tau_{si} \cdot \frac{w_s L_s / A_s L_s}{w_i L_i / A_i L_i}\right]^{\frac{\sigma-1}{\sigma}} = \left(\frac{\pi_{si}}{\pi_{ii}}\right)^{\frac{1}{\sigma}} \left[\tau_{si} \cdot \frac{w_s L_s / w_i L_i}{A_s L_s / A_i L_i}\right]^{\frac{\sigma-1}{\sigma}} \tag{3.12}$$

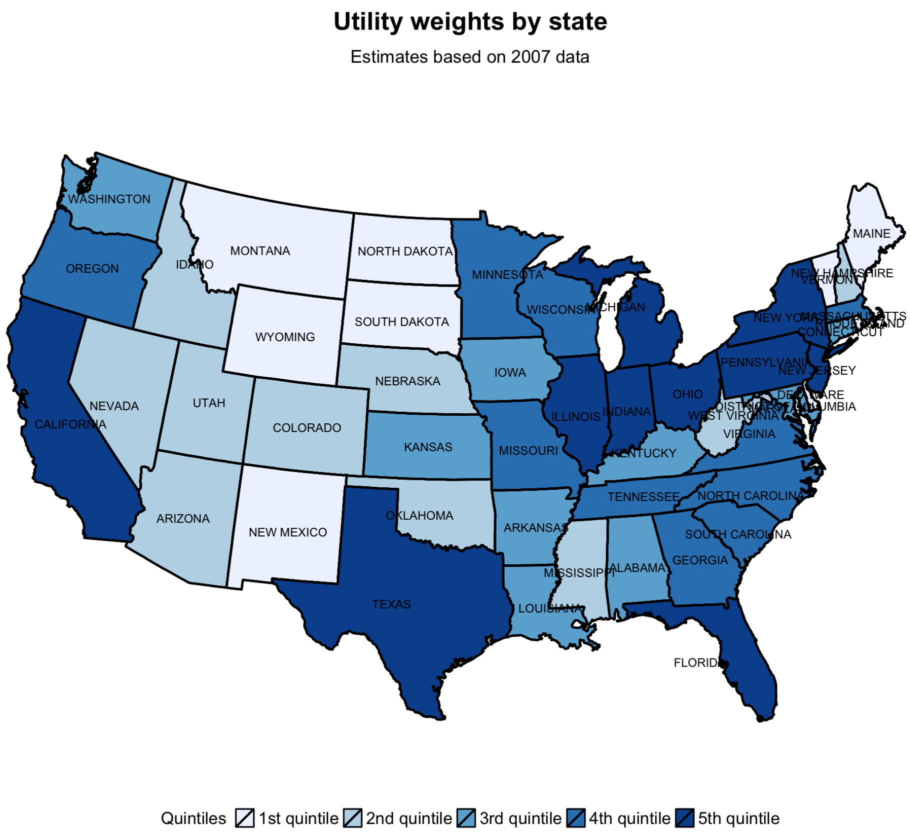


Figure 3.17: Estimates of utility weights by state

If we were estimating state-level parameters, for instance, $w_s L_s / w_i L_i$ could be proxied by the ratio of total wages from the QCEW and $A_s L_s / A_i L_i$ will be proxied by the ratio of real GDPs from the Bureau of Economic Analysis. Utility weights obtained can then be averaged across states and normalized $\sum_i a_i = 1$. As shown in Figure 3.17, this procedure results in largest utility weights assigned to bigger states like California, Texas, and Illinois while giving the lowest weight to Washington D.C. Goods produced in North Dakota and Montana by this measure have some of the lowest utility weights in the consumption preferences of a representative U.S. household.

However, in the absence of county-level data, the above procedure results in too aggregated an estimate for parameters that likely significantly vary across counties even within the same state. In particular, goods produced in metropolitan counties likely have much larger utility weights relative to more rural counties. For the empirical portion of this chapter, therefore, all counties are assumed to have equal utility weights.

Results

Estimation results generally suggest that the extent of spatial spillovers from the Bakken shale energy boom on employment in neighboring areas may be relatively modest.

The shale energy boom predictably had the strongest impact in the core Bakken counties. The counterfactual path predicted mostly flat growth in employment for the Bakken counties so that almost all of the observed growth in employment can be attributed to the energy boom. According to the estimation results, counties within 100 miles from the Bakken experienced the next largest impact. Employment growth from third quarter of 2003 to third quarter of 2014 in these counties is estimated to have been 19 percent higher relative to the counterfactual path in the absence of the oil boom. The estimated ripple effects beyond 100 miles are relatively modest, with counties in the '100 to 200 miles' and '200 to 300 miles' bands estimated to have had employment growth that is only 2 and 1 percent, respectively, higher relative to the counterfactual. The procedure predicts no impact from the energy boom on employment beyond 300 miles away from the Bakken.

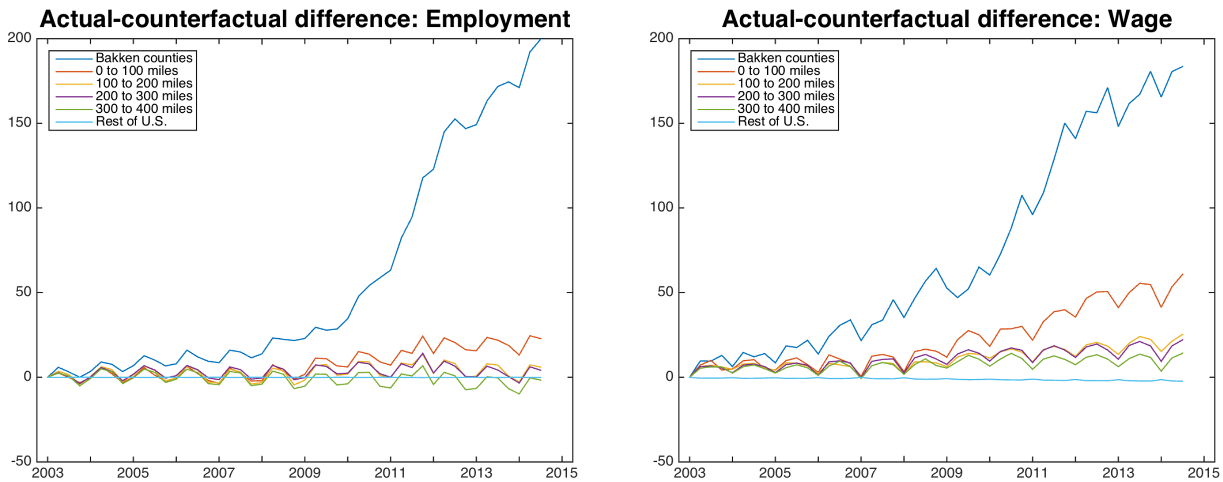


Figure 3.18: Estimated impact of the Bakken oil boom relative to counterfactual paths

Region	Actual (Data)	Counterfactual (Model)	Difference
Bakken area	187.9	2.1	185.8
0 to 100 miles	26.6	7.2	19.4
100 to 200 miles	13.5	11.6	2.0
200 to 300 miles	15.7	14.6	1.0
300 to 400 miles	11.8	16.2	-4.3
Rest of U.S.	5.9	6.0	-0.0

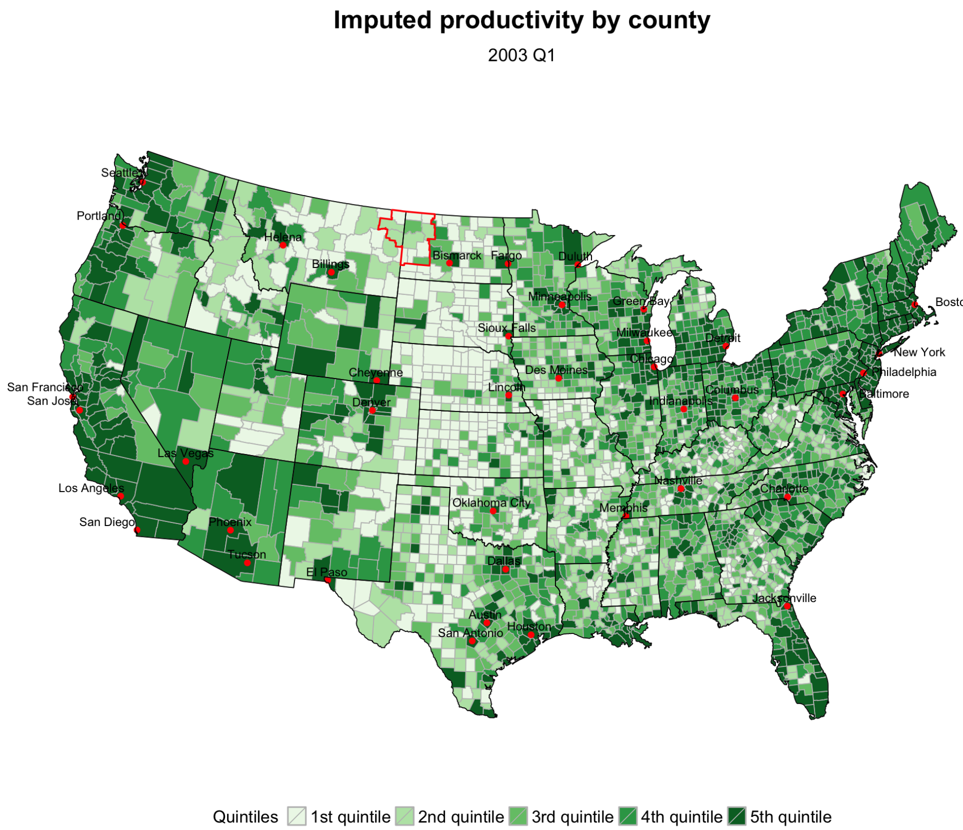
Table 3.1: Employment growth, percent change from 2003Q3 to 2014Q3

Spillover effects on wages, on the other hand, are estimated to have had a much wider impact. Growth in average weekly wages at the epicenter of the oil boom is estimated to have been 167 percent higher than the counterfactual, and the estimated spillover effects show a dissipating pattern with distance away from the epicenter. In contrast to employment spillovers, the oil boom's estimated ripple effects extend beyond the '0 to 100 miles' band, showing 6.8 percent difference between actual and counterfactual growth rates in wages even for counties in the '300 to 400 miles' band.

Region	Actual (Data)	Counterfactual (Model)	Difference
Bakken area	195.9	28.6	167.4
0 to 100 miles	81.8	31.9	49.9
100 to 200 miles	52.1	35.2	16.9
200 to 300 miles	50.2	37.0	13.2
300 to 400 miles	44.4	37.7	6.8
Rest of U.S.	34.2	35.8	-1.7

Table 3.2: Wage growth, percent change from 2003Q3 to 2014Q3

In examining the results, it is also interesting to look at the imputed productivities resulting from the estimation procedure. Maps below plot estimated productivities by county for first quarter of 2003 and second quarter of 2014, respectively. The parametrized model generally attributes higher relative productivities to major metropolitan counties. Notably, Bakken area counties are inferred to have risen in relative productivity ranking over this time period, many of them rising from the lowest quintile group to the top quintile.



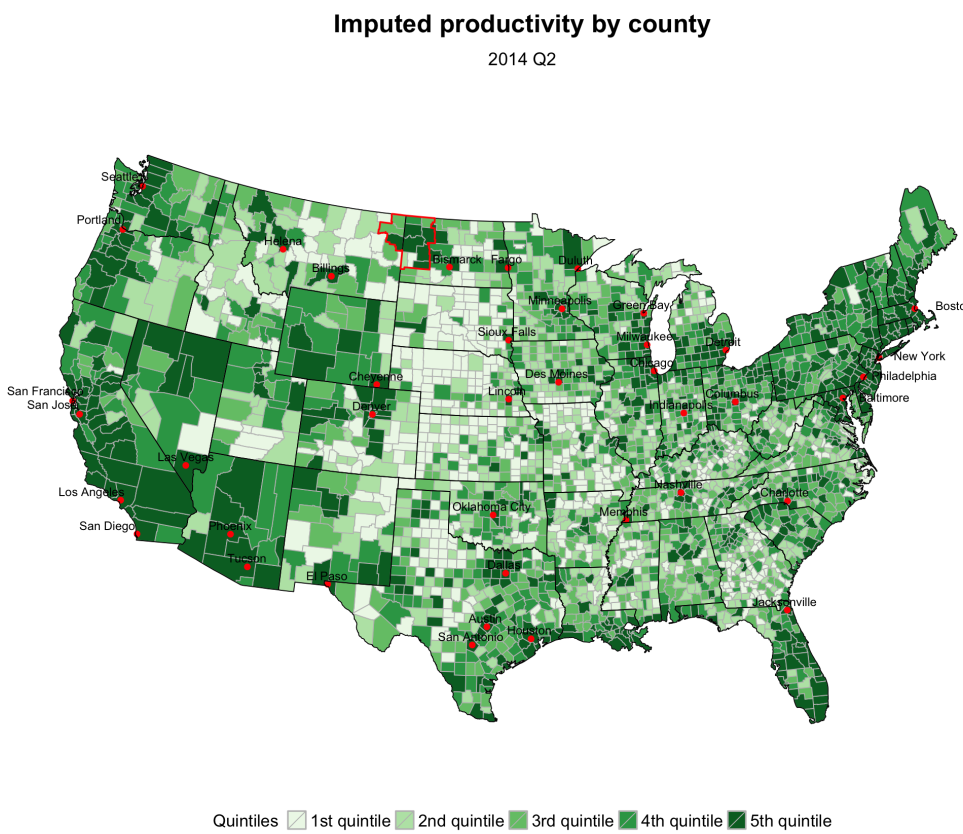


Figure 3.19: Productivities in Bakken counties are estimated to have increased from the lowest quintiles to the highest in the nation

If we chart the imputed county-level productivity growth between the two periods, Bakken area counties report highest growth rates in the nation. Moreover, counties in general areas known to have had significant shale oil and gas plays are also imputed to have had some of the highest productivity growths, according to the procedure.

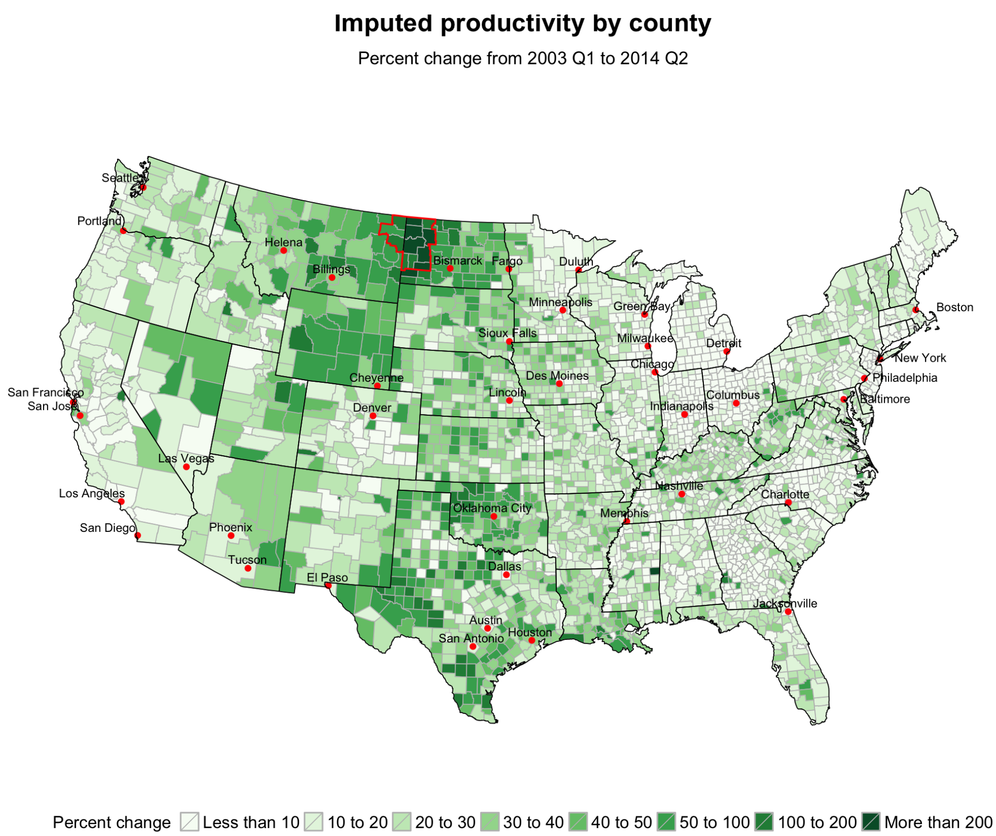


Figure 3.20: Estimated productivity growth strongest in the Bakken

Chapter 4

Economic Impact of the Oyu Tolgoi Project

Developing nations with large endowments of mineral resources often lack the technical knowledge or the financial capacity to develop these resources on their own. In such cases, strategic partnerships with established multinational companies are often seen as solutions to both problems, balancing the need for the national government to retain control over its strategic resources with the need to attract foreign investments. Mongolia is a prime example of this approach, having chosen to enter into a strategic partnership with the global mining company Rio Tinto Group to develop its Oyu Tolgoi copper-gold deposit, with the expectation that the partnership would raise productivity, increase employment and boost average incomes both in the immediate resource region and beyond. More generally, such large-scale mining projects have been pursued by the government as a leverage through which to advance its regional development objectives. This chapter uses the spatial general equilibrium model to assess the efficacy of such arrangements in achieving the government goals by evaluating the extent of spatial spillover effects on regional economies from the first-phase investments into the Oyu Tolgoi project. These early estimates based on the structural model suggest limited spillover effects on productivity and employment so far as well as possibly negative spillovers on neighboring regions' wages.

Introduction

Although developing nations rich in mineral resources have the option of growing their economies by taking advantage of their resource wealth, they often face significant constraints on their ability to put those resources to economic use.

One such constraint stems from the fact that mining is a highly capital-intensive industry. According to Schlesinger et al (2011), for instance, establishing a full copper production complex (from a mine to refinery) is estimated to require a fixed investment of about $30,000 USD per ton of copper production capacity as well as an additional 10 percent in requisite working capital. Correspondingly, a large mining complex capable of producing 100,000 tons of copper per year can easily require well over $3 billion USD in initial investments, a sum that is often beyond the means of developing nations like Mongolia, where such an investment can equal well over 25 percent of the gross national output.

Furthermore, of the total estimated investment cost reported above, construction of an open-pit mine alone accounts for about a third of capital expenditures, while an underground mine can cost up to five times as much. A concentrator plant, next on the supply chain, accounts for another third of the total investment requirement. Consequently, even investments limited to the upstream-end of the production chain still require substantial capital investments. Moreover, the investment cost hurdle may be further elevated by the need to concurrently develop necessary infrastructure in terms of establishing access to energy, water, as well as transportation links to end markets, especially if the mineral deposit is located in a remote region with little existing infrastructure.

Mining investments are also characterized by high levels of uncertainty and risk, stemming from the highly irreversible nature of capital investments, susceptibility to the price volatility of the underlying commodity, technical uncertainties related to the quality and accessibility of ore contained in the target deposit, as well as political risks. Developing nations also often lack the knowledge and expertise to implement large scale mining projects and face significant difficulties, as a result, in securing financing for such projects on their own.

The necessity of involving foreign investors can create a tension between the investor company's private profit motives and those of the government, whose objectives are typically much broader in scope and can include such long term goals as raising average incomes, promoting job creation, supporting regional development, and improving competitiveness of the domestic industry.

In such cases, strategic partnerships with well-established multinational companies are often seen as a viable solution, balancing the perceived need by the national governments to retain control over the use of strategic resources with the need to attract foreign investments. Mongolia is a prime example of this approach, having chosen to enter into a strategic partnership with the global mining company Rio Tinto Group as an investor to develop its Oyu Tolgoi copper-gold deposit.

This paper uses a spatial general equilibrium model to evaluate the economic impact of the first-phase of investments into the Oyu Tolgoi project on regional economies. These estimates based on the structural model suggest very modest effects to date on productivity and employment and possibly a negative impact on the neighboring aimags' wages.

4.1 Oyu Tolgoi Project

Project History

Mongolia is a sparsely populated nation of only 3 million people, a large fraction of whom still maintain traditional nomadic lifestyle herding animals. Prior to the onset of a major mineral exploration and mining boom in mid to late 2000s, Umnugovi, Dundgovi, and Dornogovi (administrative regions in the Gobi desert belt) have generally been the lesser populated of the 21 Mongolian aimags, especially relative to major population centers to the north around the capital city.

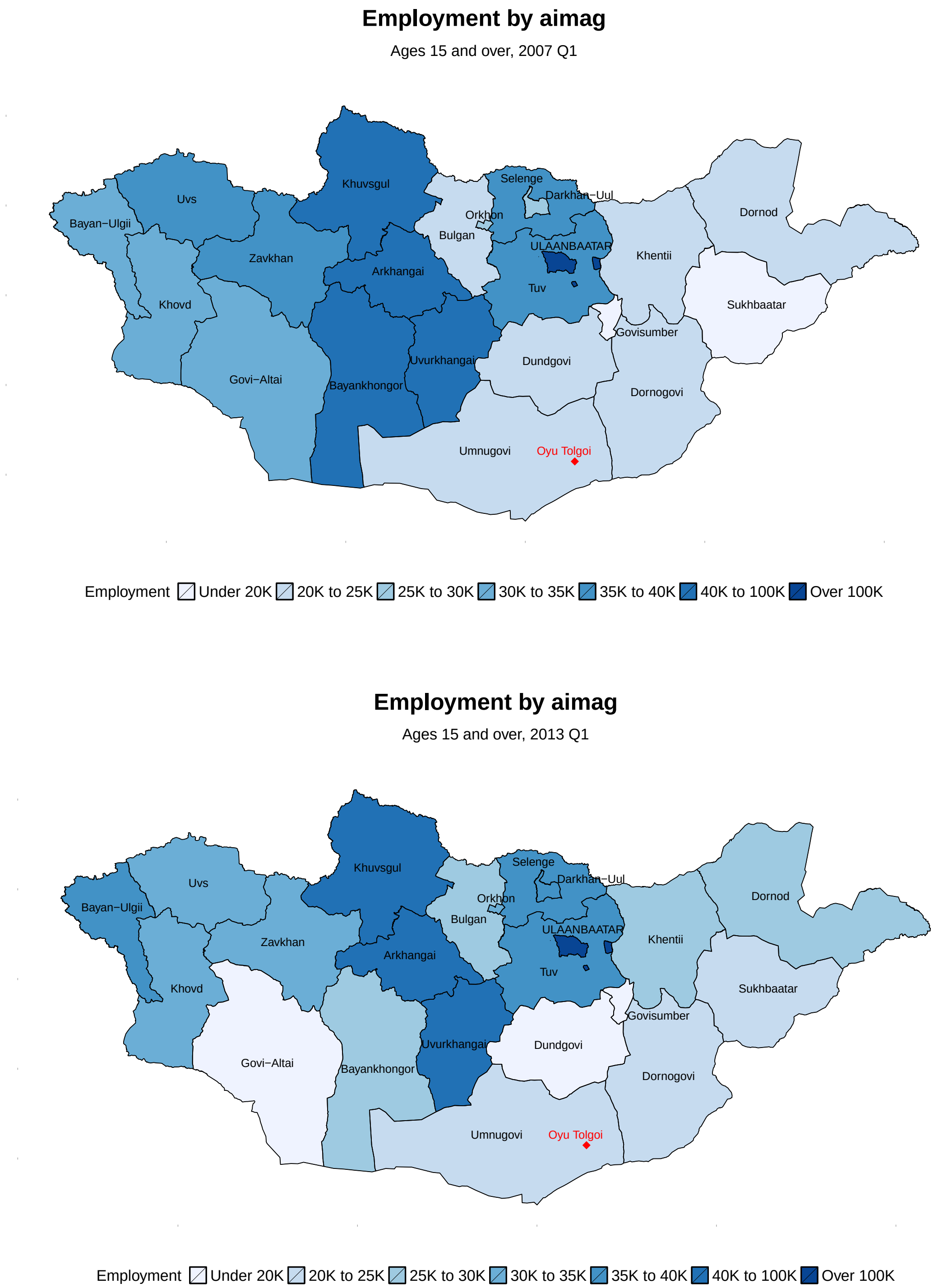


Figure 4.1: Mongolia's population and employment increasingly concentrated in central aimags close to the capital city

However, discoveries of major deposits of mineral resources such as coal, copper, and fluorspar have led to an exploration and mining boom in these aimags, culminating in the inauguration of the Oyu Tolgoi project in 2010, named after a rich copper-gold deposit located in the southeastern part of the Umnugovi aimag, just 80 kilometers north of the southern border of Mongolia with China, with estimated reserves of 21 million tons of copper and 700 tons of gold content (Turquoise Hill Resources, 2014). While the area was long before known to contain copper ores, commercially viable deposits were first discovered only relatively recently in 2001 by a Canadian exploration company Ivanhoe Mines, whose controlling share was subsequently acquired by the Rio Tinto Group.

In October of 2009, the Mongolian Government signed the Oyu Tolgoi Investment Agreement, establishing a strategic partnership with private investors to build and operate the Oyu Tolgoi mining complex. Rio Tinto is the chief stakeholder and manager in the partnership through its controlling share in the Turquoise Hill Resources Ltd that owns 66 percent of shares in the Oyu Tolgoi project, while the Mongolian government retains equity stake in the remaining 34 percent through its state-owned company Erdenes Oyu Tolgoi LLC.

Following ratification of the Agreement in early 2010, the OT project commenced its first phase of investments, starting construction of its open-pit mining complex and the concentrator plant. The first phase of the project was successfully completed in 2013, when Oyu Tolgoi started its first shipments of copper concentrate for export to China. The company reports having spent over $4.6 billion USD (Oyu Tolgoi, 2014) in Mongolia between 2010 and 2014, a significant expenditure relative to the size of the Mongolian economy, whose GDP in 2010 was measured at just over $6 billion USD.

Government Objectives

The Oyu Tolgoi Investment Agreement contains a number of provisions that indicate both explicit and implicit objectives pursued by the Government in structuring the agreement. While the primary purpose of the Agreement was to set the overall framework for the partnership, establishing the rights and obligations of each party, setting rules for taxation, dispute resolution, etc., the language of the agreement also signals three broad objectives that are being pursued by the government.

One such objective repeatedly mentioned in the agreement is the goal of regional development. Paragraph 1.7 of the Agreement, for instance, commits the investor company to make its "best effort" to promote regional development of the Southern Gobi region. As detailed in Chapter 4 of the Agreement, these efforts would include a mandatory membership in the Regional Development Council, charged with a wide range of responsibilities from preparing regional development strategies to solving urban planning and development issues, with a priority focus on residents of the Umnugovi aimag.

Another key government objective implicit in the language of the agreement is raising the productivity of the domestic mining industry, both through import of modern technologies and business practices by the investing firm as well as through training of Mongolian workers. On the technology side, Paragraph 3.11, for instance, mandates the investor to adopt a "modern mining and processing technology." Paragraph 6.20, likewise, requires the investor to "apply modern technology and procedures" to maximize efficiency of water usage by the mine. On the training aspect, Paragraph 8.12 obligates the investor to come up with a five-year "Training Strategy and Plan," which must include funding of programs aimed at improving vocational and professional skills of the Mongolian workforce, both on-the-job by the firm as well as through scholarships for studies at universities at home and abroad.

Lastly, a consistent thread visibly weaved into the Agreement is the government's push to maximize domestic employment generated by the project. This intent is strongly featured in Paragraph 8.4, setting a 90 percent minimum for the domestic share of the mine's workforce; Paragraph 8.5 that likewise establishes minimum quotas for Mongolian workers during the construction phase; and Paragraph 8.11, which commits the investing firm to ensuring that majority of its employed engineers are Mongolian citizens within a set number of years from the start of the production phase.

Now that the first phase of the project is complete, one might want to assess just how successful the Mongolian government's unorthodox method of embedding its goals into an investment agreement has been in achieving them. This chapter evaluates the government's success in pursuing the above objectives by employing the spatial general equilibrium model to estimate the economic impact on wages and employment of the Umnugovi aimag and beyond.

Spatial pattern of impact

For ease of exposition, I follow the same approach adopted in Chapter 3 and aggregate aimags into 200-kilometer (about 125 mile) bands by their geographical distance to the Oyu Tolgoi mine for reporting and analysis. This results in 6 partitions of all aimags in Mongolia, including the Ulaanbaatar capital city as a stand-alone category.

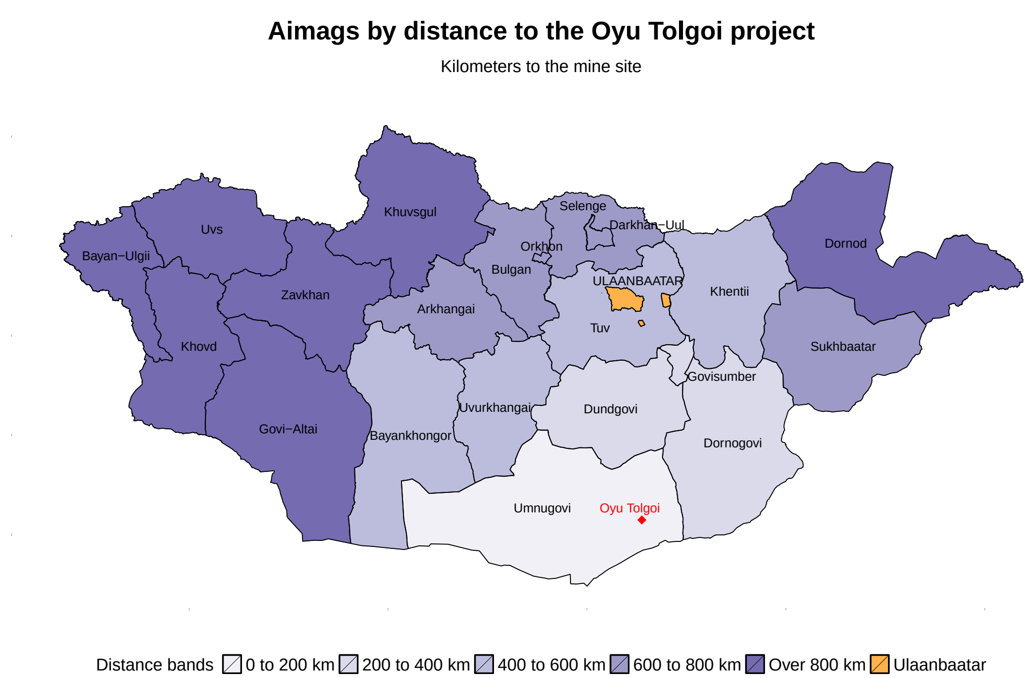


Figure 4.2: Distances are calculated using aimags' centerpoint latitude and longitude coordinates

Looking at the employment and wage data (4-quarter moving averages), aggregated by 200 kilometer bands in this manner, one can make three observations. First, the capital city Ulaanbaatar has by far dwarfed all other aimags in terms of its employment growth, despite having had the slowest average wage growth over the same period. Given its sheer weight and importance, therefore, the capital city district will be assigned its own category in the spatial analysis to follow. Second, Umnugovi - host to the Oyu Tolgoi mine and the only aimag whose centerpoint coordinate is within 200 kilometers of the mine's location - appears to have experienced only a temporary boost to its employment, which started several quarters ahead of the signing of the OT Investment Agreement and appears to have ended shortly before the official end of the first phase of the OT project.

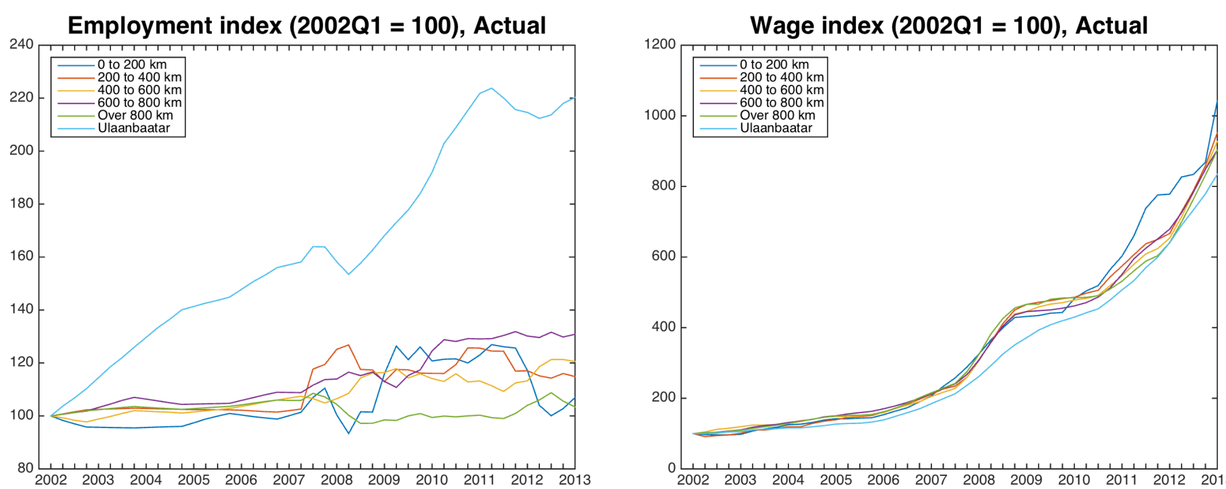


Figure 4.3: Wage growth strongest in aimags geographically closer to the Oyu Tolgoi mine

Third, the same general spatial pattern can be observed in the wage growth data as the one we've seen in the Bakken case. In particular, wage growth over the past decade has been strongest in the Umnugovi aimag, while its immediate neighbors had had the second highest wage growth, and so on, in a general hierarchy by their distance from the Oyu Tolgoi mine.

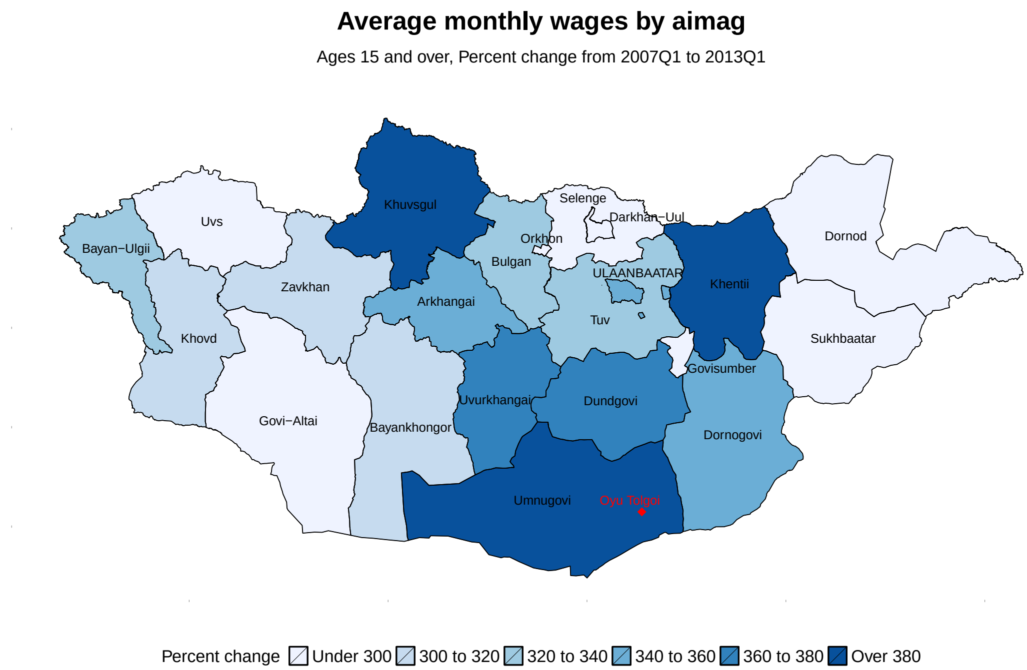


Figure 4.4: Gobi region aimags experienced faster than average wage growth during the mining boom years

As discussed in Chapter 3, however, these raw observations can be misleading given that even a mediocre wage growth could be potentially hiding a substantial impact if the counterfactual growth can be shown to be much lower.

4.2 Oyu Tolgoi Economic Impact

Data

This chapter will generally apply the same estimation procedure outlined in Chapters 2 and 3 of this thesis. For the imputation procedure, all Mongolian data, including quarterly real GDP, regional employment and wage series, are obtained for 21 aimags and capital district of Mongolia from the National Statistical Office. The analysis is done on the quarterly series from the third quarter of 2001 to the first quarter of 2013, smoothed as 4-quarter moving averages. There are three reasons why the chosen interval doesn't cover the full span of the available data, which extends to fourth quarter of 2014. The first reason has to do with limited data availability as the quarterly wage series have a long break in 2013, missing second through fourth quarter observations. Secondly, this paper's focus is specifically on the first phase of the OT project, which started in 2010 and was completed by the second half of 2013. Lastly, starting from the second half of 2013 and well through year 2014, the Mongolian government has been in the midst of a prolonged dispute with the investor Rio Tinto company, which froze all pending investments, delaying the implementation of the second phase of the project, which may confound the project's economic impact estimates for the dispute periods.

In terms of parameters used, transportation costs τ_{ij} are again assumed to be proportional to geographical distances between regions. Geographic distances are approximated using centerpoint longitude and latitude coordinates of Mongolian aimags from the shapefiles. The Earth's curvature is, as before, taken into account using haversine formula, assuming a perfectly spherical shape to the planet.

This chapter will continue to use as its baseline Armington elasticity of $\sigma = 5$, estimated earlier in Chapter 3 using U.S. Commodity Flow Survey data. To the best of my knowledge, Mongolia does not yet have the equivalent of the survey that can be used to estimate the Armington elasticity parameter, so the analysis will assume the same parameter for Mongolia. To compensate for this assumption, results will be reported for a range of plausible estimates.

To estimate the congestion parameter β specific to Mongolia, this chapter assumes the Rosen-Roback interpretation of the congestion externality. As shown in Chapter 3 of the thesis, the congestion parameter can in such cases be estimated as a function of the representative household's expenditure share on non-traded goods: $\beta = -\alpha/(1-\alpha)$. According to the National Statistical Office data, the share of household expenditures spent on rent and utilities averages at about 12 percent for a typical Mongolian household, which would imply a congestion parameter of about $\beta = -0.12/(1-0.12) = -0.14$.

The congestion parameter estimated above for Mongolia is much smaller in absolute terms compared to its U.S. counterpart estimated in Chapter 3. This is likely due to a combination of factors such as the lower density of population in Mongolia and the large share of population who live in traditional Mongolian gers, which are typically much less costly relative to more permanent dwellings.

Results

Estimation results reported in 4.1 show very small spatial spillovers from the OT project on employment in neighboring aimags, especially considering the sheer size of investments made relative to the national GDP. First phase investments into the OT mine are shown to have had the strongest impact on the Umnugovi aimag's own employment and wages, increasing them by 3 and 21 percent relative to the counterfactual, respectively. However, its fellow Gobi region aimags (200 to 400 km band) appear to have had their employment grow at best by only a fraction of a percentage point relative to the counterfactual even while having nearly no spillover effect on employment growth in other parts of the country. If anything, the overall number of jobs is estimated to have actually fallen as a result, though the decline is too small to be statistically meaningful.

Region	Actual (Data)	Counterfactual (Model)	Difference
Umnugovi	-11.5	-14.2	2.77
200 to 400 km	-1.1	-1.2	0.14
400 to 600 km	5.7	5.6	0.07
600 to 800 km	5.1	5.2	-0.11
Over 800 km	4.0	4.1	-0.06
Ulaanbaatar	14.8	14.9	-0.16

Table 4.1: Employment growth, percent change from 2010Q1 to 2013Q1

The model results, tabulated in 4.2, also suggest that the OT project may have had a negative spillover effect on average monthly wages in neighboring aimags.

Region	Actual (Data)	Counterfactual (Model)	Difference
Umnugovi	116.1	94.8	21.33
200 to 400 km	95.9	97.1	-1.20
400 to 600 km	93.9	95.1	-1.16
600 to 800 km	95.7	96.9	-1.17
Over 800 km	86.1	87.2	-1.11
Ulaanbaatar	94.2	95.4	-1.15

Table 4.2: Wage growth, percent change from 2010Q1 to 2013Q1

In Tables 4.3 and 4.4 below, I present the impact estimates (third columns in tables 4.1 and 4.2) for a range of assumed Armington elasticities. One general conclusion one can draw from these is that impact estimates for employment effects are more sensitive to the parameterization of σ while wage effects are more robust. Regardless of the choice of the Armington elasticity parameter, however, spillover effects are consistently estimated to be very small for both employment and average wages.

Region	$\sigma = 3$	$\sigma = 5$ (Base)	$\sigma = 7$	$\sigma = 9$
Umnugovi	-10.97	2.77	12.02	17.68
200 to 400 km	0.23	0.14	0.83	-0.59
400 to 600 km	0.25	0.07	0.32	-0.46
600 to 800 km	0.27	-0.11	0.09	-1.05
Over 800 km	0.24	-0.06	-2.15	1.60
Ulaanbaatar	0.31	-0.16	0.41	-1.59

Table 4.3: Estimated employment impact, percent change from 2010Q1 to 2013Q1

Region	$\sigma = 3$	$\sigma = 5$ (Base)	$\sigma = 7$	$\sigma = 9$
Umnugovi	19.18	21.33	23.13	24.89
200 to 400 km	-1.05	-1.20	-1.32	-1.41
400 to 600 km	-1.04	-1.16	-1.28	-1.26
600 to 800 km	-1.05	-1.17	-1.29	-1.38
Over 800 km	-0.99	-1.11	-1.05	-1.41
Ulaanbaatar	-1.05	-1.15	-1.28	-1.37

Table 4.4: Estimated wage impact, percent change from 2010Q1 to 2013Q1

Chapter 5

Conclusion and Discussion

Given the limitations of nonstructural empirical methods of estimating spatial spillovers of economic booms like the one observed in the Bakken region, this thesis presented an alternative method based on a computationally tractable model of a spatial general equilibrium to estimate these effects.

A common theme observed from the two applications of the model is that booms in extractive industries are largely confined to the immediate resource regions with relatively little spillover effects to neighboring areas.

In the case of the Oyu Tolgoi project, the lack of a strong impact is surprising given the sheer size of the investments made relative to the size of the economy. It is possible that more time may be needed before the full impact is reflected on official statistics. It is also possible that this study's focus on employment and wages is missing the possible economywide effects transmitted through non-labor market channels such as government revenues or direct spillover effects on other foreign investments (e.g. through a reputational channel). Significant measurement errors may also be playing a role, given the large informal sector in the country and its very mobile population.

An important omission of this approach, subject for further investigation, is the labor flows in and out of unemployment, which has also been shown to exhibit a spatially correlated pattern, particularly around the Bakken counties. Future work will need to be directed towards more robustness checks, assessing sensitivity of results to calibrated parameters of the model as well as alternative specifications for counterfactual paths.

References

[1] H. Alcott and D. Kenniston. Dutch Disease or Agglomeration? The Local Economic Effects of Natural Resource Booms in Modern America. University of Illinois at Urbana-Champaign, unpublished manuscript, 2013.

[2] T. Allen and C. Arkolakis. Trade and the Topography of the Spatial Economy. *Quarterly Journal of Economics*, 129(3):1085–1140, 2014.

[3] J.E. Anderson. A Theoretical Foundation for the Gravity Equation. *American Economic Review*, 69(1):106–116, 1979.

[4] P.S. Armington. A Theory of Demand for Products Distinguished by Place of Production. *International Monetary Fund Staff Papers*, 16(1):159–178, 1969.

[5] Baker Hughes Investor Relations. North America Rotary Rig Count: US Drilling Type 1991-2013.

[6] D. Batbold and R. Grunewald. Bakken activity: How wide is the ripple effect. *fedgazette*, Federal Reserve Bank of Minneapolis, July 2013.

[7] C.M. Brandt. Impact of the Bakken Oil Boom on Employment and Wages in North Dakota. *Undergraduate Economic Review*, 10(1), 2013.

[8] N. Brown, H. Fossum, A. Hecht, C. Dorrington, and D. McBroom. Impact of Bakken Region Oil Development on Montana's Transportation and Economy. Technical report, Montana Department of Transportation, January 2013.

[9] Bureau of Economic Analysis. Relative Price Parity.

[10] Bureau of Labor Statistics. Quarterly Census of Employment and Wages.

[11] J. Cartwright and M. Huuse. 3D seismic technology: the geological Hubble. *Basin Research*, 17(1):1–20, 2005.

[12] Census Bureau. 2010 TIGER/Line Shapefiles.

[13] Census Bureau. County-to-County Migration Flows: 2008-2012 ACS.

[14] J. Cochener. Quantifying Drilling Efficiency. Technical report, Energy Information Administration, June 2010.

[15] Department of Transportation. Commodity Flow Survey, 2007.

[16] G. Downton and A. Hendricks. New Directions in Rotary Steerable Drilling. *Journal 226*, 26:29, 1999.

[17] J. Eaton and S. Kortum. Technology, Geography, and Trade. *Econometrica*, 70(5):1741–1779, 2002.

[18] Energy Information Administration. Pad Drilling and Rig Mobility Lead to More Efficient Drilling. *Today in Energy*, Energy Information Administration, September 2012.

[19] Energy Information Administration. Shale in the United States. *Energy in Brief*, September 2014.

[20] Erdenes Oyu Tolgoi LLC. About Oyu Tolgoi Project.

[21] T. Fitzgerald. Frackonomics: Some Economics of Hydraulic Fracturing. *Case Western Reserve Law Review*, 63(4):1337–1362, 2013.

[22] G. Frobenius. Uber Matrizen aus Positiven Elementen. *Reports of the Prussian Academy of Sciences*, pages 514–518, 1909.

[23] S.B. Gawirth, K.R. Marra, T.A. Cook, R.R. Charpentier, D.L. Gautier, D.K. Higley, T.R. Klett, M.D. Lewan, P.G. Lillis, C.J. Schenk, M.E. Tennyson, and K.J. Whidden. Assessment of undiscovered oil resources in the Bakken and Three Forks formations, Williston Basin Province, Montana, North Dakota, and South Dakota, 2013. *U.S. Geological Survey Fact Sheet*, 20133013:4, 2013.

[24] R. Grunewald and D. Batbold. Booming sales in North Dakota. *fedgazette*, Federal Reserve Bank of Minneapolis, April 2013.

[25] Internal Revenue Service. SOI Tax Stats - County-to-County Migration Data Files.

[26] J. Marchand. The Distributional Impacts of an Energy Boom in Western Canada. Working Paper No. 2013-13, Vancouver School of Economics, 2014.

[27] Mineral Resource Authority of Mongolia. Geology and Mining Sector Report, 2014.

[28] E. Moretti. Local Labor Markets. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 4b. North-Holland, 2010.

[29] National Statistical Office of Mongolia. Employed Aged 15 and Over, by sex, regions, aimags and the Capital.

[30] National Statistical Office of Mongolia. Gross Domestic Product, by quarter, by expenditure approach.

[31] National Statistical Office of Mongolia. Monthly Average Wages and Salaries, by aimags and the Capital, sex, quarter.

[32] North Dakota Industrial Commission. Monthly Production Reports.

[33] Oyu Tolgoi. Oyu Tolgoi Scorecard - Monitoring Our Perfomance, 2014.

[34] O. Perron. Zur Theorie der Matrizen. *Mathematische Annalen*, 64:248–263, 1907.

[35] J. Roback. Wages, Rents and the Quality of Life. *Journal of Political Economy*, 90(6):1257–1278, 1982.

[36] S. Rosen. Wage-based indexes of urban quality of life. In P.N. Miezkowski and M.R. Straszheim, editors, *Current Issues in Urban Economics*, pages 74–104. Johns Hopkins University Press, Baltimore, MD, 1979.

[37] I. Runge. Mining Economics. Presented at the Discover Mongolia 2012 International Mining Conference and Investors Forum, Ulaanbaatar, Mongolia, 2012.

[38] M. Schlesinger, M. King, K. Sole, and W. Davenport. Costs of Copper Production. In M. Schlesinger, editor, *Extractive Metallurgy of Copper*, chapter 22. Elsevier, United Kingdom, 5th edition, 2011.

[39] T. Tunstall and J. Oyakawa. Economic Impact of the Eagle Ford Shale. Technical report, Center for Community and Business Research, University of Texas at San Antonio, September 2014.

[40] Turquoise Hill Resources Ltd. Oyu Tolgoi - 2014 Technical Report. Technical report, Turquoise Hill Resources Ltd., October 2014.

Appendix A

Two-Region Example

Consider an illustrative example with $S = \{1, 2\}$. Then (2.13) can be written as:

$$\begin{bmatrix} a_1^\sigma A_1^{\sigma-1} & a_2^\sigma A_2^{\sigma-1}\tau^{1-\sigma} \\ a_1^\sigma A_1^{\sigma-1}\tau^{1-\sigma} & a_2^\sigma A_2^{\sigma-1} \end{bmatrix} \cdot \begin{bmatrix} w_1^{1-\sigma} \\ w_2^{1-\sigma} \end{bmatrix} = \lambda \cdot \begin{bmatrix} w_1^{1-\sigma} \\ w_2^{1-\sigma} \end{bmatrix} \tag{A.1}$$

where $\tau_{12} = \tau_{21} = \tau$ is the transportation cost of shipping goods between the two regions. If $\tau > 1$, then the two eigenvalues for the 2x2 square matrix in (A.1) will be given by:

$$\lambda = \frac{a_1^\sigma A_1^{\sigma-1} + a_2^\sigma A_2^{\sigma-1}}{2} \pm \frac{\sqrt{\left(a_1^\sigma A_1^{\sigma-1} - a_2^\sigma A_2^{\sigma-1}\right)^2 + 4 \cdot a_1^\sigma a_2^\sigma \cdot A_1^{\sigma-1} A_2^{\sigma-1} \cdot \tau^{2(1-\sigma)}}}{2}$$

According to the Perron-Frobenius Theorem, only one (the larger) eigenvalue of the two is associated with a strictly positive eigenvector, so that welfare $W^{\sigma-1}$ equals to the eigenvalue λ with the plus sign. Note that welfare (i.e. real wages) would then be strictly increasing in productivities and decreasing in τ, which is reasonable given that higher productivities would mean more output produced and larger iceberg transportation costs would mean more output lost due to "melting" in transit. We can then use

$$0 = \begin{bmatrix} a_1^\sigma A_1^{\sigma-1} - \lambda & a_2^\sigma A_2^{\sigma-1}\tau^{1-\sigma} \\ a_1^\sigma A_1^{\sigma-1}\tau^{1-\sigma} & a_2^\sigma A_2^{\sigma-1} - \lambda \end{bmatrix} \cdot \begin{bmatrix} w_1^{1-\sigma} \\ w_2^{1-\sigma} \end{bmatrix}$$

to solve for relative wages:

$$\left[\frac{w_2}{w_1}\right]^{\sigma-1} = \frac{\lambda - a_2^\sigma A_2^{\sigma-1}}{a_1^\sigma A_1^{\sigma-1}\tau^{1-\sigma}} = \tau^{\sigma-1}\cdot\frac{1-\hat{a}}{2} + \sqrt{\left(\tau^{\sigma-1}\cdot\frac{1-\hat{a}}{2}\right)^2 + \hat{a}} \tag{A.2}$$

where $\hat{a} \equiv a_2^\sigma A_2^{\sigma-1}/a_1^\sigma A_1^{\sigma-1}$. Equation (A.2) shows three properties of the equilibrium relative wages. First, note that in the absence of transportation costs ($\tau = 1$), the law of one price holds with $P_1 = P_2$. Correspondingly, welfare equalization would also imply $w_1 = w_2$, consistent with (A.2) where the wedge between regional wages appears only if the transportation cost τ is greater than 1.

Second, equation (A.2) also implies that the wage wedge is bounded by the transportation cost τ. To see this, note that

$$\lim_{\hat{a}\to\infty}\frac{w_2}{w_1} = \frac{1}{\tau} \qquad \text{and} \qquad \lim_{\hat{a}\to 0}\frac{w_2}{w_1} = \tau \tag{A.3}$$

so that $1/\tau < w_2/w_1 < \tau$, i.e. the greater the distance, the wider is the band between which relative wages can diverge away from 1 in either direction.

Finally, note that in the symmetric case (i.e. $\hat{a} = 1$), prices would also equalize. Moreover, taking the derivative of (A.2) with respect to $\hat{a}$, one can observe that:

$$\frac{\partial (w_2/w_1)^{\sigma-1}}{\partial \hat{a}} = -\frac{\tau^{\sigma-1}}{2} + \frac{1}{2\sqrt{\ldots}}\cdot\left(1 - \tau^{2(\sigma-1)}\cdot\frac{1-\hat{a}}{2}\right) = \frac{1}{2\sqrt{\ldots}}\cdot\left[1 - \left(\tau\cdot\frac{w_2}{w_1}\right)^{\sigma-1}\right] < 0 \tag{A.4}$$

since $w_2/w_1 > 1/\tau$, which means in the two-region version of the model, a positive productivity shock in a given region induces lower nominal wages in that region relative to that of its neighbor.

We can likewise rewrite (2.17) for the two-region case as follows:

$$\begin{bmatrix} a_1^\sigma A_1^{\sigma-1} & \tau^{1-\sigma} a_1^\sigma A_1^{\sigma-1}\cdot(w_2/w_1)^\sigma \\ \tau^{1-\sigma} a_2^\sigma A_2^{\sigma-1}\cdot(w_1/w_2)^\sigma & a_2^\sigma A_2^{\sigma-1} \end{bmatrix}\cdot\begin{bmatrix} L_1 \\ L_2 \end{bmatrix} = \phi\cdot\begin{bmatrix} L_1 \\ L_2 \end{bmatrix} \tag{A.5}$$

Note that consistent with (2.13) and (2.17), the square matrix in (A.5) will have the exact same eigenvalues as in (A.1), yielding a consistent real wage estimate $\phi = \lambda = W^{\sigma-1}$. Relative employment then will be characterized by:

$$\frac{L_2}{L_1} = \frac{\tau^{1-\sigma} a_2^{\sigma} A_2^{\sigma-1}}{\phi - a_2^{\sigma} A_2^{\sigma-1}} \cdot \frac{w_1^{\sigma}}{w_2^{\sigma}} \;=\; \left[\tau^{\sigma-1} \cdot \frac{1/\hat{a} - 1}{2} + \sqrt{\left(\tau^{\sigma-1} \cdot \frac{1/\hat{a} - 1}{2}\right)^2 + 1/\hat{a}}\,\right]^{-1} \cdot \frac{w_1^{\sigma}}{w_2^{\sigma}} \;=\; \hat{a} \cdot \frac{w_1}{w_2} \tag{A.6}$$

Just as with wages, the two regions in the symmetric case ($\hat{a} = 1$) will share the workers equally, i.e. $L_1 = L_2$. Unlike the wages, however, relative employment shares are not bounded. Note that the righthand side of (A.6) approaches infinity as the relative productivity parameter $\hat{a}$ grows larger since we know relative wages are bounded, as shown in (A.3). Consequently,

$$\lim_{\hat{a} \to 0} \frac{L_2}{L_1} = 0 \qquad \text{and} \qquad \lim_{\hat{a} \to \infty} \frac{L_2}{L_1} = \infty \tag{A.7}$$

Note that (A.6) also implies that $w_2 L_2 / w_1 L_1$ is directly proportional to relative productivities, i.e. the more productive region will have a higher income.